arXiv:math-ph/0701066v1 27 Jan 2007

HARMONIC ANALYSIS OF ITERATED FUNCTION SYSTEMS WITH OVERLAP

PALLE E. T. JORGENSEN, KERI A. KORNELSON, AND KAREN L. SHUMAN

Dedicated to the memory of Thomas P. Branson

Abstract. An iterated function system (IFS) is a system of contractive mappings $\tau_i \colon Y \to Y$, $i = 1, \ldots, N$ (finite) where Y is a complete metric space. Every such IFS has a unique (up to scale) equilibrium measure (also called the Hutchinson measure μ), and we study the Hilbert space $L^2(\mu)$.

In this paper we extend previous work on IFSs without overlap. Our method involves systems of operators generalizing the more familiar Cuntz relations from operator algebra theory, and from subband filter operators in signal processing.

These Cuntz-like operator systems were used in recent papers on wavelets analysis by Baggett, Jorgensen, Merrill and Packer, where they serve as a first step to generating wavelet bases of Parseval type (alias normalized tight frames), i.e., wavelet bases with redundancy.

Similarly, it was shown in work by Dutkay and Jorgensen that the iterative operator approach works well for generating wavelets on fractals from IFSs without overlap. But so far the more general and more difficult case of essential overlap has resisted previous attempts at a harmonic analysis, and explicit basis constructions in particular.

The operators generating the appropriate Cuntz relations are composition operators, e.g., $F_i \colon f \to f \circ \tau_i$ where (τ_i) is the given IFS. If the particular IFS is essentially non-overlapping, it is relatively easy to compute the adjoint operators $S_i = F_i^*$, and the S_i operators will be isometries in $L^2(\mu)$ with orthogonal ranges. For the case of essential overlap, we can use the extra terms entering in the computation of the operators F_i^* as a "measure" of the essential overlap for the particular IFS we study. Here the adjoint operators F_i^* refers to the Hilbert space $L^2(\mu)$ where μ is the equilibrium measure μ for the given IFS (τ_i).

Contents

2000 *Mathematics Subject Classification.* 33C50, 42C15, 46E22, 47B32.

Key words and phrases. tight frames, Parseval frames, Bessel sequences, Hilbert space, polar decomposition, attractor, IFS, equilibrium measure, composition operator.

This material is based upon work supported by the U.S. National Science Foundation under grants DMS-0457581 and DMS-0503990.

1. Introduction: IFSs and operator theory

In this paper we explore an operator theoretic approach to self similarity and fractals which is common to problems involving both iterated function systems (IFSs), and an aspect of quantum communication.

Each of the two areas involves a finite set of operations: in the case of IFSs they are geometric, and in the quantum case, they involve channels of Hilbert spaces and associated operator systems. The particular aspects of IFSs we have in mind are studied in [DJ06a]; and the relevant results from quantum communication in [KLPL06] and [Kri05]. We begin the Introduction with some background on IFSs, and we motivate our present operator theoretic approach. The operator theory and its applications are then taken up more systematically in section 2 below.

It was proved recently that a class of non-linear fractals is amenable to a computational harmonic analysis. While this work is motivated by applications, it is significant in its own right; for example these fractals do not carry the structure of groups, and so they do not come with a Haar measure. Nonetheless, they have natural equilibrium measures μ which serve as substitutes. Specifically, these fractals are known as iterated function systems (IFS) indicating the recursive nature of their construction. But the analysis so far has been limited to IFSs without essential overlap (see Definition 3.9). The IFS fractals X are built by passing to a certain limit in a scaling iteration, and the "overlap" here refers to the parts of smaller scale-level. We identify "overlap" for a given X relative to the equilibrium measure μ for X. While there is prior work on some isolated cases of overlap, the cases that are understood are those of negligible overlap. In this paper we identify and study the harmonic analysis of IFSs with "substantial overlap."

An iterated function system (IFS) is a system of contractive mappings $\tau_i\colon Y \to Y$, $i = 1, \ldots, N$ (finite) where Y is a complete metric space. Every such IFS has a unique (up to scale) equilibrium measure (also called the Hutchinson measure μ), and we study the Hilbert space $L^2(\mu)$.

In this paper we extend previous work on IFSs without overlap. Our method involves systems of operators generalizing the more familiar Cuntz relations from operator algebra theory, and from subband filter operators in signal processing. Before turning to the details, below we outline briefly the operator-theoretic approach to IFSs.

For each N, there is a simple Cuntz C^*-algebra on generators and relations, and its representations offer a useful harmonic analysis of general IFSs, but there is a crucial difference between IFSs without overlap and those with essential overlap (this will be made precise in Section 2). The operators generating the appropriate Cuntz relations [Cun77] are composition operators, e.g., $F_i\colon f \to f \circ \tau_i$ where (τ_i) is the given IFS. If the particular IFS is essentially non-overlapping, it is relatively easy to compute the adjoint operators $S_i = F_i^*$, and the S_i operators will be isometries in $L^2(\mu)$ with orthogonal ranges. In a way, for the more difficult case of

essential overlap, we can use the extra terms entering into the computation of the adjoint operators F_i^* as a "measure" of the essential overlap for the particular IFS we study. Here the adjoint operators F_i^* refer to the Hilbert space $L^2(\mu)$ where μ is carefully chosen. When the IFS is given, there are special adapted measures μ. We will be using the equilibrium measure μ for the given IFS (τ_i), and by [Hut81], this μ contains much essential information about the IFS, even in the classical cases of IFSs coming from number theory.

So far, earlier work in the general area was restricted to IFSs without overlap. For example, Cuntz-like operator systems were used in recent papers on wavelets analysis by Baggett, Jorgensen, Merrill and Packer [BJMP04, BJMP05, BJMP06], where they serve as a first step to generating wavelet bases of Parseval type (alias normalized tight frames), i.e., wavelet bases with redundancy.

Similarly, it was shown in work by Dutkay and Jorgensen [DJ06d] that the iterative operator approach works well for generating wavelets on fractals from IFSs without overlap. But so far the more general and more difficult case of essential overlap has resisted previous attempts at a harmonic analysis, and explicit basis constructions in particular.

In this paper (Section 4) we show that certain combinatorial cycles from a symbolic encoding [DJ06b] of our IFSs yield an attractive computational analysis of the "IFS-overlap." Some results (e.g., Corollary 4.11) are stated only for a model example, but as indicated the idea with cycle-counting works in general.

While the modern study of iterated function systems (IFSs) has roots in classical problems from number theory, infinite products [Kol77], and more generally from probability and harmonic analysis (see, e.g., [FLP94]), the subject took a more systematic turn in 1981 with the influential paper [Hut81] by Hutchinson. Since then, IFSs have found uses in geometry (e.g., [Bar06, BHS05]), in infinite network problems (e.g., [GRS01]), in wavelets (e.g., [BJMP05, BJMP06, Jor05]), and in dynamics and operator theory [Kaw05, Jor06, DJ06c, DJ06b, DJ05].

But in operator algebras, the subject has its own independent start with the paper by Cuntz [Cun77]. In operator-algebra theory, a class of representations of what have become known as the Cuntz, or the Cuntz–Krieger, algebras is ideally suited for the study of iterative processes (including IFSs) in mathematics and in mathematical physics. Perhaps the paper [CK80] started this trend, but it was continued in various guises in the following papers: see [Pop89] and later related papers by Popescu, [BV05], [JP96], [JK03], and [Jor04], among others. The Cuntz and the related C^*-algebras have become ubiquitous tools in the study of iterative systems, or in the analysis of fractals or of graphs. One reason for this is that there is an unexpected connection to signal and image processing; see [Jor06]. They are algebras on generators and relations; and often, as is the case in the present paper, the generators may be identified as substitution operators [Jor04]. However, these substitution operators and their adjoints (often called transfer operators) are of independent interest; see, e.g., [JP96, Sin05, KS93, Kwa04].

Since the Cuntz and related C^*-algebras typically are naturally generated by a finite set of concrete operators, it has become customary to study such finite sets as matrices with operator entries; for example, a class of vectors with operator entries are studied in [Arv04] as row contractions. In fact, it is convenient to distinguish between row vectors with operator entries, and the corresponding columns of operators: one is the adjoint of the other. In this paper, our study of IFSs with overlap

leads naturally to a family of column isometries, where the operator entries in our column vectors are substitution operators built from the maps that define the IFS under consideration.

While there is already a rich literature on IFSs without overlap, the more difficult case of overlap has received relatively less attention; see, however, [Bar06, FLP94, Sol95, Sol98]. The point of view of the latter three papers is generalized number expansions in the form of random variables: Real numbers are expanded in a basis which is a fraction, although the "digits" are bits; with infinite strings of bits identified in a Bernoulli probability space. It turns out the distribution of the resulting random variables is governed by the measures which arise as a special cases of Hutchinson's analysis [Hut81] of IFSs with overlap. See also [HR00]. However, concrete results about these measures have been elusive. For example, it is proved in [Sol95] that the measures for expansions in a basis corresponding to IFSs with overlap and given by a scaling parameter are known to be absolutely continuous for a.e. value of the parameter.

In this paper we consider a rather general class of IFSs with overlap. We show that they can be understood in terms of the spectral theory of Cuntz-like column isometries. Moreover we show that our column isometries yield exact representations of the Cuntz relations precisely when the IFS has overlap of measure zero, where the measure is an equilibrium measure μ of the Hutchinson type.

There is a separate development related to the study of Fourier bases in the Hilbert spaces $L^2(\mu)$ for μ an IFS equilibrium measure; see [JP98, Str00, SU00, LW00, Jor06, DJ06b]. We address some of these issues below, but it turns out that exact Fourier bases for the case of IFS with overlap are harder to come by. Further, in [DJ06b], the coauthors noted that Fourier bases in $L^2(\mu)$ can only exist if the equilibrium measure μ has equal weights, and so in the present paper, we make this assumption.

Recent references [DJ05, DJ06c, DJ06d, DJ06b, Jor06] deal with wavelet constructions on non-linear attractors X, such as arise from systems of branched mappings, e.g., finite affine systems (affine IFSs), or Julia sets [Bea91] generated by branches of the inverses of given rational mappings of one complex variable. These X come with associated equilibrium measures μ. However, due to geometric obstructions, it is typically not possible for the corresponding $L^2(X,\mu)$ to carry orthonormal bases (ONBs) of complex exponentials, i.e., Fourier bases. We showed instead in [DJ06d] that wavelet bases may be constructed in ambient $L^2(\mu)$ Hilbert spaces via multiresolutions and representations of Cuntz-like operator relations in $L^2(X,\mu)$ as in Definitions 2.1 and 2.3 in the next section.

2. Multivariable operator theory

In quantum communication (the study of (quantum) error-correction codes), certain algebras of operators and completely positive mappings form the starting point; see especially the papers [KLPL06] and [Kri05]. They take the form of a finite number of channels of Hilbert space operators F_i which are assumed to satisfy certain compatibility conditions. The essential one is that the operators from a partition of unity, or rather a partition of the identity operator I in the chosen Hilbert space. Here (Definition 2.1) we call such a system (F_i) a column isometry. An extreme case of this is when a certain Cuntz relation (Definition 2.3) is satisfied by (F_i). Referring back to our IFS application, the extreme case of the

operator relations turn out to correspond to the limiting case of non-overlap, i.e., to the case when our IFSs have no essential overlap (Definition 3.9).

In this section we outline some uses of ideas from multivariable operator theory (see, e.g., [Arv04]) in iterated function systems (IFS) with emphasis on *IFSs with overlap*. The tools we use are *column isometries* and systems of *composition operators*.

Definition 2.1. Let $\mathcal{H}$ be a complex Hilbert space, and let $N \in \mathbb{N}$, $N \geq 2$. A system $(F_1, \dots, F_N)$ of bounded operators in $\mathcal{H}$ is said to be a *column isometry* if the mapping

$$\mathbb{F} \colon \mathcal{H} \longrightarrow \begin{pmatrix} \mathcal{H} \\ \oplus \\ \vdots \\ \oplus \\ \mathcal{H} \end{pmatrix} : \xi \longmapsto \begin{pmatrix} F_1 \xi \\ \vdots \\ F_N \xi \end{pmatrix} \tag{2.1}$$

is isometric. Here we write the N-fold orthogonal sum of $\mathcal{H}$ in column form, but we will also use the shorter notation $\mathcal{H}_N$. As a Hilbert space, it is the same as $\mathcal{H} \oplus \cdots \oplus \mathcal{H}$, but for clarity it is convenient to identify the *adjoint operator* $\mathbb{F}^*$ as a row

$$\mathbb{F}^* \colon \underbrace{\mathcal{H} \oplus \cdots \oplus \mathcal{H}}_{N \text{ times}} \longrightarrow \mathcal{H} : (\xi_1, \dots, \xi_N) \longrightarrow \sum_{i=1}^{N} F_i^* \xi_i. \tag{2.2}$$

The inner product in $\mathcal{H}_N$ is $\sum_{i=1}^{N} \langle \xi_1 \mid \eta_i \rangle$, and relative to the respective inner products on $\mathcal{H}_N$ and on $\mathcal{H}$, we have

$$\left\langle \underbrace{\mathbb{F}}_{\substack{\text{column} \\ \text{operator}}} \xi \;\middle|\; \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_N \end{pmatrix} \right\rangle_{\mathcal{H}_N} = \left\langle \xi \;\middle|\; \underbrace{\mathbb{F}^*}_{\substack{\text{row} \\ \text{operator}}} \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_N \end{pmatrix} \right\rangle_{\mathcal{H}}. \tag{2.3}$$

Remark 2.2. In the definition of a column isometry, it is stated that $\mathbb{F}$ in (2.1) is *isometric* $\mathcal{H} \to \mathcal{H}_N$, but it is not necessarily onto $\mathcal{H}_N$. This means that in general the matrix operator $\mathbb{F}\mathbb{F}^* \colon \mathcal{H}_N \to \mathcal{H}_N$ is a proper *projection*. By this we mean that the block matrix

$$\mathbb{F}\mathbb{F}^* = \left(F_i F_j^*\right)_{i,j=1}^{N} \tag{2.4}$$

satisfies the following system of identities:

$$\sum_{k=1}^{N} \left(F_i F_k^*\right)\left(F_k F_j^*\right) = F_i F_j^*, \qquad 1 \leq i, j \leq N. \tag{2.5}$$

Note that $\mathbb{F}\mathbb{F}^* = I_{\mathcal{H}_N}$ if and only if

$$F_i F_j^* = \delta_{i,j} I, \qquad 1 \leq i, j \leq N. \tag{2.6}$$

Definition 2.3. A column isometry $\mathbb{F}$ satisfies $\mathbb{F}\mathbb{F}^* = I_{\mathcal{H}_N}$ if and only if it defines a representation of the Cuntz algebra $\mathcal{O}_N$. In that case, the operators $S_i := F_i^*$ are isometries in $\mathcal{H}$ with orthogonal ranges, and

$$\sum_{i=1}^{N} S_i S_i^* = I_{\mathcal{H}}. \tag{2.7}$$

Remark 2.4. The distinction between the operator relations in Definitions 2.1 and 2.3 is much more than a technicality: Definition 2.3 is the more restrictive. Because of the orthogonality axiom in Definition 2.3, it is easy to see that if a Hilbert space $\mathcal{H}$ carries a nonzero representation of the Cuntz relations (Definition 2.3) (see [Cun77]), then $\mathcal{H}$ must be infinite-dimensional, reflecting the infinitely iterated and orthogonal subdivision of projections, a hallmark of fractals.

In contrast, the condition of Definition 2.1, or equivalently $\sum_{i=1}^{N} F_i^* F_i = I_{\mathcal{H}}$, may easily be realized when the dimension of the Hilbert space $\mathcal{H}$ is finite. In fact such representations are used in quantum computation; see, e.g., [LS05, Theorem 2] and [Kri05].

Definition 2.5. Let $(X, \mathcal{B}, \mu)$ be a finite measure space, i.e., X is a set, $\mathcal{B}$ is a sigma-algebra of subsets in X, and μ is a finite positive measure defined on $\mathcal{B}$. We will assume that $\mathcal{B}$ is complete, and the Hilbert space $L^2(X, \mathcal{B}, \mu)$ will be denoted $L^2(\mu)$ for short. If ν is a second measure, also defined on $\mathcal{B}$, we say that $\nu \ll \mu$ (relative absolute continuity) if

$$S \in \mathcal{B}, \qquad \mu(S) = 0 \implies \nu(S) = 0. \tag{2.8}$$

In that case, the corresponding Radon–Nikodym derivative will be denoted $\varphi := \frac{d\nu}{d\mu}$; i.e., we have $\varphi \in L^1(\mu)$, and

$$\nu(S) = \int_S \varphi(x)\, d\mu(x) \qquad \text{for all } S \in \mathcal{B}. \tag{2.9}$$

An endomorphism $\tau\colon X \to X$ is said to be *measurable* if

$$S \in \mathcal{B} \implies \tau^{-1}(S) \in \mathcal{B}, \tag{2.10}$$

where $\tau^{-1}(S) = \{\, x \in X \mid \tau(x) \in S \,\}$. In that case, a measure $\mu \circ \tau^{-1}$ is *induced* on $\mathcal{B}$, and given by

$$\left(\mu \circ \tau^{-1}\right)(S) := \mu\left(\tau^{-1}(S)\right), \qquad S \in \mathcal{B}. \tag{2.11}$$

Remark 2.6. Every measurable endomorphism $\tau\colon X \to X$, referring to $(X, \mathcal{B}, \mu)$, induces a *composition operator*

$$C_\tau\colon L^2(\mu) \ni f \longmapsto f \circ \tau \in L^2(\mu), \tag{2.12}$$

and it can be shown that C_τ is a *bounded* operator in the Hilbert space $L^2(\mu)$ if and only if $\mu \circ \tau^{-1} \ll \mu$ with Radon–Nikodym derivative in L^∞. Moreover, if $\mu \circ \tau^{-1} \ll \mu$, set $\varphi := \frac{d\mu \circ \tau^{-1}}{d\mu}$.

Definition 2.7. A measurable endomorphism $\tau\colon X \to X$ is said to be of *finite type* if there are a *finite partition* $E_1, \dots, E_k$ of $\tau(X)$ and measurable mappings $\sigma_i\colon E_i \to X$, $i = 1, \dots, k$ such that

$$\sigma_i \circ \tau|_{E_i} = \mathrm{id}_{E_i}, \qquad 1 \le i \le k. \tag{2.13}$$

Lemma 2.8. *Let $(X, \mathcal{B}, \mu)$ be as above, and let $\tau\colon X \to X$ be a measurable endomorphism of finite type. Suppose $\mu \circ \tau^{-1} \ll \mu$, and set $\varphi := \frac{d\mu \circ \tau^{-1}}{d\mu}$.*

- *(a) Then φ is supported (μ-a.e.) on $\tau(X)$.*
- *(b) Moreover, if $E_1, \dots, E_k$ is a (non-overlapping) partition as in* (2.13)*, then the adjoint operator C_τ^* of the composition operator* (2.13) *is given by the formula*

$$C_\tau^* f|_{E_i} = \varphi\,(f \circ \sigma_i), \qquad i = 1, \dots, k,\ f \in L^2(\mu); \tag{2.14}$$

i.e., for $x \in E_i$, $\left(C_\tau^* f\right)(x) = \varphi(x)\, f\left(\sigma_i(x)\right)$.

Proof. We will begin by assuming $\mu \circ \tau^{-1} \ll \mu$. At the end of the resulting computation, it will then be clear that the reasoning is in fact reversible. The composition operator C_τ is well defined as in (2.13), referring to the Hilbert space $L^2(\mu)$ with its usual inner product

$$\langle f_1 \mid f_2 \rangle_\mu := \int_X \overline{f_1}\, f_2\, d\mu. \tag{2.15}$$

For the adjoint operator C_τ^*, we have

$$\begin{aligned}
\langle C_\tau^* f_1 \mid f_2 \rangle_\mu &= \langle f_1 \mid C_\tau f_2 \rangle_\mu = \langle f_1 \mid f_2 \circ \tau \rangle_\mu \\
&= \int_X \overline{f_1}\, (f_2 \circ \tau)\, d\mu = \sum_{i=1}^{k} \int_{\tau^{-1}(E_i)} \overline{f_1}\, (f_2 \circ \tau)\, d\mu \\
&= \sum_{i=1}^{k} \int_{\tau^{-1}(E_i)} \left(\overline{f_1} \circ \sigma_i \circ \tau\right) (f_2 \circ \tau)\, d\mu \\
&= \sum_{i=1}^{k} \int_{E_i} \left(\overline{f_1} \circ \sigma_i\right) f_2\, d\mu \circ \tau^{-1} \\
&= \sum_{i=1}^{k} \int_{E_i} \left(\overline{f_1} \circ \sigma_i\right) f_2\, \varphi\, d\mu
\end{aligned}$$

for all $f_1, f_2 \in L^2(\mu)$. The desired formula (2.14) for the adjoint operator C_τ^* follows. □

Definition 2.9. Let $(X, \mathcal{B}, \mu)$ be a finite measure space. A system $\tau_1, \dots, \tau_N$ of measurable endomorphisms is said to be an *iterated function system* (*IFS*). Let $p_i \in [0,1]$ be given such that $\sum_{i=1}^{N} p_i = 1$. If

$$\sum_{i=1}^{N} p_i \mu \circ \tau_i^{-1} = \mu \tag{2.16}$$

we say that the measure μ is a *p-equilibrium* measure. Motivated by earlier work on the harmonic analysis of equilibrium measures (see, e.g., [JP98] and [DJ06b]), we will here restrict attention to the case of equal weights, i.e., assume that $p_1 = \cdots = p_N = 1/N$; and we shall then refer to μ as an *equilibrium measure*, with the understanding that $p_i = 1/N$. In general, equilibrium measures might not exist; and even if they do, they may not be unique.

Proposition 2.10. *Let $(X, \mathcal{B}, \mu)$ be a finite measure space, and let $\tau_1, \dots, \tau_N$ be measurable endomorphisms. Then some μ is an equilibrium measure if and only if the associated linear operator*

$$\mathbb{F}_\tau \colon L^2(\mu) \longrightarrow \begin{pmatrix} L^2(\mu) \\ \oplus \\ \vdots \\ \oplus \\ L^2(\mu) \end{pmatrix} : f \longmapsto \frac{1}{\sqrt{N}} \begin{pmatrix} f \circ \tau_1 \\ \vdots \\ f \circ \tau_N \end{pmatrix} \tag{2.17}$$

is isometric, i.e., if and only if the individual operators

$$F_i : f \longmapsto \frac{1}{\sqrt{N}} f \circ \tau_i \qquad \text{in } L^2(\mu) \tag{2.18}$$

define a column isometry.

Proof. Using polarization for the inner product $\langle \cdot \mid \cdot \rangle_\mu$ in (2.15), we first note that μ is an equilibrium measure if and only if

$$\frac{1}{N} \sum_{i=1}^{N} \int_X |f|^2 \circ \tau_i \, d\mu = \int_X |f|^2 \, d\mu \qquad (=: \|f\|^2_{L^2(\mu)}) \tag{2.19}$$

holds for all $f \in L^2(\mu)$.

The terms on the left-hand side in (2.19) are $\frac{1}{N} \int_X |f|^2 \circ \tau_i \, d\mu = \|F_i f\|^2_{L^2(\mu)}$, so (2.19) is equivalent to

$$\left\langle f \,\middle|\, \sum_{i=1}^{N} F_i^* F_i f \right\rangle_\mu = \sum_{i=1}^{N} \langle F_i f \mid F_i f \rangle = \sum_{i=1}^{N} \|F_i f\|^2_{L^2(\mu)} = \|f\|^2_{L^2(\mu)},$$

which in turn is the desired operator identity

$$\sum_{i=1}^{N} F_i^* F_i = I_{L^2(\mu)} \tag{2.20}$$

that defines $\mathbb{F}$ as a column isometry. □

Our main result in Sections 3 and 4 will yield a formula for the Radon–Nikodym derivatives $\varphi_i := \frac{d\mu \circ \tau_i^{-1}}{d\mu}$; see Remark 3.6 for details.

3. Contractive iterated function systems

The study of contractive iterated function systems (contractive IFSs) was initiated in a systematic form by Hutchinson in [Hut81], but was also used in harmonic analysis before 1981.

Suppose $N \in \mathbb{N}$, $N \geq 2$, is given, and suppose some mappings $\tau_i : Y \to Y$, $i = 1, \dots, N$ are contractive in some complete metric space (Y, d): then we say that $(\tau_i)_{i=1}^N$ is a *contractive IFS*. The contractivity entails constants $c_1, \dots, c_N$, $c_i \in (0,1)$, such that

$$d(\tau_i(x), \tau_i(y)) \leq c_i d(x,y), \qquad x, y \in Y. \tag{3.1}$$

Hutchinson's first theorem [Hut81] states that there is then a unique compact subset $X \subset Y$ such that

$$X = \bigcup_{i=1}^{N} \tau_i(X). \tag{3.2}$$

This set X is called the *attractor* for the system.

If weights $(p_i)_{i=1}^N$ are given as in Definition 2.9, Hutchinson's second theorem states that there is a unique (up to scale) positive finite (nonzero) Borel measure μ on Y such that

$$\mu = \sum_{i=1}^{N} p_i \mu \circ \tau_i^{-1}. \tag{3.3}$$

Moreover, if $p_i > 0$ for all i, then the support of μ is the compact set X (fractal) in (3.2), the attractor.

For many purposes, it is convenient to normalize the measure μ in (3.3) such that $\mu(X) = \mu(Y) = 1$. Because of applications to harmonic analysis [DJ06b], we will also restrict the weights (p_i) in (3.3) such that $p_i = 1/N$.

Definition 3.1. For Borel measures ν on Y (some given complete metric space), set

$$T\nu := \frac{1}{N}\sum_{i=1}^{N} \nu \circ \tau_i^{-1}. \tag{3.4}$$

Set

$$\mathrm{Lip}_1(Y) = \{\, f : Y \to \mathbb{R},\ |f(x) - f(y)| \le d(x,y) \,\},$$

and

$$d_1(\nu_1, \nu_2) := \sup\left\{ \int f\, d\nu_1 - \int f\, d\nu_2 \;\middle|\; f \in \mathrm{Lip}_1(Y) \right\}. \tag{3.5}$$

Then it is easy to see [Hut81] that d_1 is a metric and that the probability measures form a complete metric space with respect to d_1.

Moreover, regardless of the choice of the initial measure ν (some probability measure on Y), the limit

$$\lim_{n\to\infty} T^n \nu = \mu \qquad \text{(in the metric } d_1\text{)} \tag{3.6}$$

exists, where μ is the equilibrium measure from (3.3), $p_i = 1/N$, and where the convergence in (3.6) is relative to the metric d_1 from (3.5).

Example 3.2. Let $\mathbb{Z}_N := \{0, 1, \ldots, N-1\}$, and set $\Omega = \mathbb{Z}_N \times \mathbb{Z}_N \times \cdots = (\mathbb{Z}_N)^{\mathbb{N}}$, i.e., the infinite Cartesian product. We equip Ω with its usual Tychonoff topology, and its usual metric. Points in Ω are denoted $\omega = (\omega_1\omega_2\ldots)$, $\omega_i \in \mathbb{Z}_N$, $i = 1, 2, \ldots$. For $n \in \mathbb{N}$, set $\omega|n = (\omega_1\omega_2\ldots\omega_n)$. Further, we shall need the one-sided shifts

$$\sigma_j(\omega_1\omega_2\ldots) = (j\omega_1\omega_2\ldots), \qquad j \in \mathbb{Z}_N,\ \omega \in \Omega, \tag{3.7}$$

and

$$\sigma(\omega_1\omega_2\omega_3\ldots) = (\omega_2\omega_3\ldots), \qquad \omega \in \Omega. \tag{3.8}$$

If $(\tau_i)_{i=1}^N$ is a contractive IFS with attractor X, it is easy to see that for each $\omega \in \Omega$, the intersection

$$\bigcap_{n=1}^{\infty} \tau_{\omega|n}(X) \tag{3.9}$$

is a singleton. Here we use the notation

$$\tau_{\omega|n} := \tau_{\omega_1}\tau_{\omega_2}\cdots\tau_{\omega_n}. \tag{3.10}$$

If $\omega \in \Omega$ is given, let $\pi(\omega)$ be the (unique) point in the intersection (3.9). The mapping $\pi\colon \Omega \to X$ is called the *encoding*.

Lemma 3.3. *Let $(\tau_i)_{i=1}^N$ and X be as above, i.e., assumed contractive. Then the coding mapping $\pi\colon \Omega \to X$ from* (3.9) *is continuous, and we have*

$$\pi \circ \sigma_j = \tau_j \circ \pi, \qquad j = 1, \ldots, N \quad \text{or} \quad j \in \mathbb{Z}_N. \tag{3.11}$$

Proof. The continuity is clear from the definitions. We verify (3.11): Let $\omega = (\omega_1\omega_2\ldots) \in \Omega$. Then

$$\begin{aligned}(\pi \circ \sigma_j)(\omega) = \pi(j\omega_1\omega_2\ldots) &= \bigcap_n \tau_j\tau_{\omega_1}\cdots\tau_{\omega_n}(X)\\ &= \tau_j\left(\bigcap_n \tau_{\omega_1}\tau_{\omega_2}\cdots\tau_{\omega_n}(X)\right)\\ &= \tau_j(\pi(\omega)) = (\tau_j\circ\pi)(\omega),\end{aligned}$$

where we used contractivity of the mappings τ_j. □

Corollary 3.4. *It follows from Hutchinson's theorem* [Hut81] *that for each* $(p_1,\ldots,p_N)$ *with* $\sum_1^N p_i = 1$, $p_i \geq 0$, *there is a unique probability measure* $P_{(p)}$ *on* Ω *such that*

$$P_{(p)} = \sum_{i=1}^N p_i P_{(p)} \circ \sigma_i^{-1}. \tag{3.12}$$

When $p_i = 1/N$, *this measure is called* the Bernoulli measure *on* Ω. *The measures* $P_{(p)}$ *are also familiar infinite-product measures, considered first by Kolmogorov* [Kol77].

Corollary 3.5. *Let* $(\tau_i)_{i=1}^N$ *be a contractive IFS, and let* (p_i) *be given such that* $\sum_{i=1}^N p_i = 1$, $p_i \geq 0$. *Let* $\pi\colon \Omega \to X$ *be the corresponding endoding mapping. Then the equilibrium measure* $\mu_{(p)}$ *for* (τ_i) *is*

$$\mu_{(p)} = P_{(p)} \circ \pi^{-1}, \tag{3.13}$$

i.e., for Borel subsets $E \subset X$ *we have*

$$\mu_{(p)}(E) = P_{(p)}\left(\pi^{-1}(E)\right), \tag{3.14}$$

where

$$\pi^{-1}(E) = \{\,\omega \in \Omega \mid \pi(\omega) \in E\,\}. \tag{3.15}$$

Proof. Let (τ_i), (p), Ω, and $P_{(p)}$ be as described. The following computation uses (3.11) in the form

$$\sigma_j^{-1}\pi^{-1} = \pi^{-1}\tau_j^{-1}, \qquad j \in \mathbb{Z}_N.$$

The conclusion will follow from the uniqueness part of Hutchinson's theorem if we check that $P_{(p)} \circ \pi^{-1}$ satisfies the identity (2.16).

Let E be a Borel set $\subset X$. Then

$$\begin{aligned}\left(P_{(p)} \circ \pi^{-1}\right)(E) &= P_{(p)}\left(\pi^{-1}(E)\right)\\ &= \sum_{i=1}^N p_i P_{(p)}\left(\sigma_i^{-1}\left(\pi^{-1}(E)\right)\right)\\ &= \sum_{i=1}^N p_i P_{(p)}\left(\pi^{-1}\left(\tau_i^{-1}(E)\right)\right)\\ &= \sum_{i=1}^N p_i \left(P_{(p)} \circ \pi^{-1}\right) \circ \tau_i^{-1}(E).\end{aligned}$$

Since $P_{(p)} \circ \pi^{-1}$ is a probability measure, it follows that $\mu_{(p)} := P_{(p)} \circ \pi^{-1}$ is the unique solution to (2.16) given by the normalization $\mu_{(p)}(X) = 1$. □

Remark 3.6. It is immediate from the definitions that all the measures μ solving (3.3) for the case of positive weights $p_i > 0$ will satisfy the relative absolute continuity condition

$$\mu \circ \tau_i^{-1} \ll \mu, \tag{3.16}$$

but we shall be especially interested in the case of equal weights $p_i = 1/N$, and in this case we set

$$\varphi_i := \frac{d\mu \circ \tau_i^{-1}}{d\mu}. \tag{3.17}$$

Example 3.7. Let $N = 2$, $\mathbb{Z}_2 = \{0, 1\}$, $\Omega = (\mathbb{Z}_2)^{\mathbb{N}}$, $p_0 = p_1 = 1/2$, and let $\lambda \in (0, 1)$ be given. Consider the two contractive maps

$$\tau_0(x) = \lambda x \quad \text{and} \quad \tau_1 x = \lambda(x + 1) \qquad \text{in } \mathbb{R}.$$

For the attractor X_λ, we have the following three cases:

Case 1: $\lambda \in (0, 1/2)$:

$$X_\lambda \text{ is a fractal of fractal dimension } D_\lambda = -\log 2/\log\lambda. \tag{3.18}$$

Case 2: $\lambda = 1/2$:

$$X_{1/2} = [\,0, 1\,], \text{ and } \mu\ (= \mu_{1/2}) \text{ is the restriction of Lebesgue measure.} \tag{3.19}$$

Case 3: $\lambda \in (1/2, 1)$:

$$X_\lambda = \left[0, \frac{\lambda}{1-\lambda}\right]; \text{ overlap.} \tag{3.20}$$

In the next section we will study these measures μ_λ (Case 3) in detail.

In all three cases, the encoding mapping is

$$\pi(\omega_1\omega_2\dots) = \sum_{i=1}^{\infty} \omega_i\lambda^i = \omega_1\lambda + \omega_2\lambda^2 + \cdots, \qquad \omega_i \in \mathbb{Z}_2 = \{0, 1\}. \tag{3.21}$$

The measure μ_λ is determined by the following approximation: If E is a Borel subset of X_λ, then

$$\mu_\lambda(E) = \lim_{n\to\infty} 2^{-n}\,\#\left\{\,(\omega_1\omega_2\dots\omega_n) \mid \omega_1\lambda + \omega_2\lambda^2 + \cdots + \omega_n\lambda^n \in E\,\right\}. \tag{3.22}$$

Proof. The only one of the stated conclusions which is not clear from Corollary 3.5 is the limit in (3.22). But (3.22) is an application of (3.6) to the case when the initial measure ν is δ_0 (the Dirac measure), i.e.,

$$\delta_0(E) = \begin{cases} 1 & \text{if } 0 \in E, \\ 0 & \text{if } 0 \notin E. \end{cases}$$

Note that

$$\begin{aligned} T^n\delta_0 &= 2^{-n} \sum_{\omega_1\omega_2\dots\omega_n} \delta_0 \circ \tau_{\omega_n}^{-1}\tau_{\omega_{n-1}}^{-1} \circ \cdots \circ \tau_{\omega_1}^{-1} \\ &= 2^{-n} \sum_{\omega_1\omega_2\dots\omega_n} \delta_{\omega_1\lambda+\omega_2\lambda^2+\cdots+\omega_n\lambda^n}. \end{aligned}$$

The summation is over $\underbrace{\{0, 1\} \times \cdots \times \{0, 1\}}_{n \text{ times}}$ for each $n = 1, 2, \dots$. □

Remark 3.8. The three cases in (3.18)–(3.20) are different in one important respect, *overlap*; and we turn to this in the next section.

Summary: OVERLAP:

$$\begin{aligned}
&\text{CASE 1: } \lambda \in (0,1/2): && \tau_0(X_\lambda) \cap \tau_1(X_\lambda) = \varnothing. \\
&\text{CASE 2: } \lambda = 1/2: && \tau_0([0,1]) \cap \tau_1([0,1]) = \left\{\frac{1}{2}\right\}. \\
&\text{CASE 3: } \lambda \in (1/2,1): && \tau_0(X_\lambda) \cap \tau_1(X_\lambda) = \left[\lambda, \frac{\lambda^2}{1-\lambda}\right].
\end{aligned}$$

Definition 3.9. Let $(X, \mathcal{B}, \mu)$ be a finite measure space, and let $(\tau_i)_{i=1}^N$ be a finite system of measurable endomorphisms, $\tau_i \colon X \to X$, $i = 1, \dots, N$; and suppose μ is some normalized equilibrium measure. We then say that the system has *essential overlap* if

$$\sum\sum_{i \neq j} \mu(\tau_i(X) \cap \tau_j(X)) > 0. \tag{3.23}$$

In Section 6, we will prove the following:

Theorem 3.10. *Let $(X, \mathcal{B}, \mu)$ and $(\tau_i)_{i=1}^N$ be as in Definition* 3.9*; in particular we assume that μ is some (τ_i)-equilibrium measure. We assume further that each τ_i is of finite type.*

Let $\mathcal{H} = L^2(\mu)$,

$$F_i \colon f \longmapsto \frac{1}{\sqrt{N}} f \circ \tau_i,$$

and let $\mathbb{F} = (F_i)$ be the corresponding column isometry

Then $\mathbb{F}$ maps onto $\bigoplus_1^N L^2(\mu)$ if and only if (τ_i) has zero μ-essential overlap.

Proof. To better understand the geometric significance of this theorem, before giving the proof, we work out a particular family of examples (Sections 4 and 5 below) in 1D and in 2D.

The intricate geometric features of IFSs can be understood nicely by specializing the particular affine transformations making up the IFS to have a single scale number (which we call λ). In the 1D examples, we will have two affine transformations (4.1), and in our 2D examples, three (5.3). (These special 1D examples also go under the name infinite Bernoulli convolutions. Here we will use them primarily for illustrating the operator theory behind Theorem 3.10.)

As shown in Section 5 below, in passing from 1D to 2D, the possible geometries of the IFS-recursions increase; for example, new fractions and new gaps may appear simultaneously at each iteration step. Specifically, (i) fractal (i.e., repeated gaps) and (ii) "essential overlap" co-exist in the 2D examples. Here "essential overlap" (Definition 3.9) is of course defined in terms of the Hutchinson measure μ $(= \mu_\lambda)$.

Our examples will be followed by a proof of Theorem 3.10 in section 6. In this concluding section we further show (Theorem 6.3) that every IFS with essential overlap (Definition 3.9) has a canonical and minimal dilation to one with non-overlap. $\square$

Remark 3.11. Note that the conclusion of the theorem states that the operators $S_i := F_i^*$ define a representation of the Cuntz C^*-algebra $\mathcal{O}_N$ if and only if the

system has non-essential overlap, i.e., if and only if $\mu\left(\tau_i\left(X\right)\cap\tau_j\left(X\right)\right)=0$ for all $i\neq j$.

4. Operator theory of essential overlap

The one-dimensional example in the previous section (Example 3.7) helps to clarify the notion of essential overlap. Recall that in this example, λ is chosen such that $1/2<\lambda<1$, and the two endomorphisms are

$$\left\{\begin{aligned}\tau_0\left(x\right)&=\lambda x,\\ \tau_1\left(x\right)&=\lambda\left(x+1\right)\end{aligned}\right\}\quad\text{and}\quad\mu=\frac{1}{2}\left(\mu\circ\tau_0^{-1}+\mu\circ\tau_1^{-1}\right)\tag{4.1}$$

with attractor $X_\lambda=\operatorname{supp}\left(\mu\right)=\left[\,0,\lambda/\left(1-\lambda\right)\right]$. But the probability measure μ $(=\mu_\lambda)$ on this interval is not the restriction of Lebesgue measure on $\mathbb{R}$. In fact, μ_λ is difficult to compute explicitly and there appears not to be a closed formula for $\mu_\lambda\left(\left[\,\lambda,\lambda^2/\left(1-\lambda\right)\right]\right)$, although the Lebesgue measure of the overlap is $\frac{\lambda(2\lambda-1)}{1-\lambda}$. It is known, for example, that μ_λ is Lebesgue-absolutely continuous for a.a. λ; see [Sol95].

From Example 3.7 it also follows that μ_λ is the distribution of the random variable $\pi_\lambda\left(\omega\right)=\sum_{i=1}^{\infty}\omega_i\lambda^i$, $\omega_i\in\{0,1\}$, or $\omega=\left(\omega_1\omega_2\dots\right)\in\Omega=\{0,1\}^{\mathbb{N}}$. Specifically,

$$P_{1/2}\left(\{\,\omega\mid\pi_\lambda\left(\omega\right)\leq x\,\}\right)=\int_{-\infty}^{x}d\mu_\lambda\left(y\right)=\mu_\lambda\left(\left[0,x\right]\right)=:F_\lambda\left(x\right).\tag{4.2}$$

Recall from Section 2 that $\operatorname{supp}\left(\mu_\lambda\right)=X_\lambda=\left[\,0,\lambda/\left(1-\lambda\right)\right].$

While explicit analytic properties (like bounded variation) satisfied by the cumulative distribution functions F_λ (for specific values of λ) aren't well understood, some geometric properties may be generated from the following proposition. It gives an intrinsic scaling identity (4.8) for F_λ, and a recursive algorithm (4.9) for its computation.

First note that by (4.2), for every λ, the function $x\to F(x)=F_\lambda(x)$ is monotone nondecreasing.

Moreover since the Hutchinson measure μ_λ is supported in the interval X_λ, F_λ is constantly 0 on the negative real line, and it is 1 on the infinite half line to the right of the bounded interval $\left[\,0,\lambda/(1-\lambda)\right]$.

Of course, F need not be strictly increasing in the interval. But since it is monotone, by Lebesgue's theorem, it is differentiable almost everywhere (for a.e. x, taken with respect to Lebesgue measure). However, a deep theorem [Sol95] states that, as a measure, μ_λ has L^2 density for a.e. λ in the open interval $(1/2,1)$; so in particular for these values of λ, the Stieltjes measure $\mu_\lambda=dF_\lambda$ is relatively absolutely continuous with respect to Lebesgue measure.

4.1. **Symmetry.** The next result (needed later!) shows that with the choice of equal weights $\frac{1}{2}$–$\frac{1}{2}$ in (4.1), the random variable $\pi_\lambda\colon\Omega\to\left[\,0,\lambda/\left(1-\lambda\right)\right]$ from (3.21) is symmetric around the midpoint

$$\frac{\lambda}{2\left(1-\lambda\right)}=\int_0^{\frac{\lambda}{1-\lambda}}x\,d\mu_\lambda\left(x\right).\tag{4.3}$$

The symmetry property which is made precise in the next lemma reflects that the x-values are corners of a hypercube: Repeated folding with shrinking intervals each with excess length.

Lemma 4.1 (Symmetry)**.** *Let* $\lambda \in (1/2, 1)$*, and let* $\mu = \mu_\lambda$ *and* $P_{1/2}$ *be as described in* (4.1) *and* (4.2)*. Recall that* $X_\lambda = \operatorname{supp}(\mu_\lambda)$ *is the closed interval* $[\,0, b(\lambda)\,]$ *where* $b(\lambda) := \lambda/(1-\lambda)$*.*

Then for all $x \leq \frac{1}{2} b(\lambda)$ *the two tail-ends of the distribution are the same, i.e.,*

$$P_{1/2}(\{\,\omega \mid \pi_\lambda(\omega) < x\,\}) = P_{1/2}(\{\,\omega \mid \pi_\lambda(\omega) > b(\lambda) - x\,\}). \tag{4.4}$$

Proof. We will prove (4.4) in the equivalent form

$$P_{1/2}(\{\,\omega \mid \pi_\lambda(\omega) \geq x\,\}) = P_{1/2}(\{\,\omega \mid \pi_\lambda(\omega) \leq b(\lambda) - x\,\}). \tag{4.5}$$

The two events we specify on the left-hand side and the right-hand side in (4.5) are described by the respective conditions

$$\sum_{i=1}^{\infty} \omega_i \lambda^i \geq x \tag{4.6}$$

and

$$\sum_{i=1}^{\infty} (1 - \omega_i)\, \lambda^i \geq x. \tag{4.7}$$

But since the "fair-coin" Bernoulli measure $P_{1/2}$ on Ω was chosen, the two sequences of independent random variables

$$\omega_1,\ \omega_2,\ \omega_3,\ \ldots$$

and

$$1-\omega_1,\ 1-\omega_2,\ 1-\omega_3,\ \ldots$$

are equi-distributed. Hence the numbers on the two sides in (4.5) are the same.

The formula (4.3) for the mean follows directly from (4.1) as follows: Set $M_1 = M_1(\lambda) = \int x\, d\mu_\lambda(x)$. Then by (4.1),

$$\begin{aligned} M_1 &= \frac{1}{2}\left(\int (\lambda x)\ d\mu_\lambda(x) + \int \lambda(x+1)\ d\mu_\lambda(x) \right) \\ &= \lambda M_1 + \frac{\lambda}{2}. \end{aligned}$$

Hence $M_1 = \frac{\lambda}{2(1-\lambda)} = \frac{1}{2} b(\lambda)$ as claimed. □

Remark 4.2. Actually every "cascade approximant" (see Proposition 4.6 below) is symmetric in a similar way around its midpoint $(1/2)\left(\lambda + \lambda^2 + \cdots + \lambda^n\right)$.

Remark 4.3. The argument at the end of the proof of the lemma extends to yield a formula for all the moments

$$M_n(\lambda) = \int x^n\, d\mu_\lambda(x),$$

beginning with

$$M_2(\lambda) = \frac{\lambda^2}{2(1-\lambda)^2 \cdot (1+\lambda)} = \frac{2M_1(\lambda)^2}{(1+\lambda)}$$

and

$$M_3(\lambda) = \frac{\lambda^3 \cdot \left(\lambda + 2 \cdot \left(1 - \lambda^2 + \lambda^3\right)\right)}{4 \cdot (1-\lambda^3)(1-\lambda^2)(1-\lambda)}.$$

The formula for $n > 1$ is recursive, and can be worked out from the binomial distribution.

There is a considerable amount of recent work (see, e.g., [CF05]) on moments in operator theory, and it would be interesting to explore the operator-theoretic significance of our present "overlap-moments."

Remark 4.4. By the argument of the previous remark, the moment-generating function

$$\hat{\mu}_\lambda(t) := \int e^{itx}\, d\mu_\lambda(x)$$

can be shown to have the following infinite-product expansion:

$$\hat{\mu}_\lambda(t) = e^{itM_1(\lambda)} \prod_{n=1}^{\infty} \cos\left(\frac{t\lambda^n}{2}\right).$$

Using Wiener's test on this, it follows that none of the measures μ_λ have atoms, i.e., that $\mu_\lambda(\{x\}) = 0$ for all x, and for all $\lambda \in [1/2, 1)$.

Remark 4.5 (Figure generation, hypercubes, and symmetry, by Brian Treadway)**.** The x-values in the figures where steps occur in F are generated using a recursive "outer sum" construction:

$$\begin{aligned} &\{0\} \to \{\{0\},\{\lambda\}\} \to \{\{\{0\},\{\lambda^2\}\},\{\{\lambda\},\{\lambda+\lambda^2\}\}\} \\ &\to \{\{\{\{0\},\{\lambda^3\}\},\{\{\lambda^2\},\{\lambda^2+\lambda^3\}\}\},\{\{\{\lambda\},\{\lambda+\lambda^3\}\},\{\{\lambda+\lambda^2\},\{\lambda+\lambda^2+\lambda^3\}\}\}\} \\ &\to \cdots. \end{aligned}$$

This nested list of the 2^n sums has the structure of the coding map—in fact, if it is expressed with "tensor indices" instead of nested braces, the indices are just our ω's. The hypercube idea comes from thinking of each ω_i as a coordinate in an orthogonal direction. Opposite corners then have x-values that sum to $(\lambda + \lambda^2 + \cdots + \lambda^n)$, so the two tails may be matched point by point. Hence the x-values are corners of a hypercube, and hence symmetry.

To make the plots of $F(x)$, the nested list above is "flattened" to a single list and sorted numerically. The *Mathematica* program that generates the sorted list is

```
Sort[Flatten[Nest[Outer[Plus, {0, l}, l #]&, {0}, n]],
        OrderedQ[{N[#1], N[#2]}]&],
```

where "`l`" stands for λ and "`n`" is the level of iteration in the cascade algorithm.

Aside from Erdős's theorem [Erd40] about the case of λ equal to the golden ratio, where μ_λ is singular, precious little is known about μ_λ for other specific λ values, for example when λ is rational.

Two interesting values of λ are $\lambda = (\sqrt{5}-1)/2$ (Fig. 2) and $\lambda = 3/4$ (Fig. 3). The cases are qualitatively different both with regard to absolute continuity and overlap and with respect to Fourier bases; see [DJ06b].

In the known results [DJ06b] for which $L^2(\mu_\lambda)$ has a Fourier orthonormal basis (ONB), λ is rational. By a Fourier basis, we mean an ONB in $L^2(\mu_\lambda)$ consisting of complex exponentials.

Because of [Erd40], $\lambda = (\sqrt{5}-1)/2$ is likely to have its F_λ a little less "differentiable" than the F_λ for $\lambda = 3/4$.

The following details will work as an iterative and cascading algorithm for $F = F_\lambda$ when λ is fixed. Our algorithm is initialized so as to take advantage of (3.22) above. It is illustrated in the figures; see especially Figure 1.

Proposition 4.6. *Let λ be given in the open interval $(1/2, 1)$, and set $F = F_\lambda$ as a function on $\mathbb{R}$. Conclusions:*

(a) Then the function $x \to F(x)$ in (4.2) satisfies

$$F(x) = \frac{1}{2}\left(F\left(\frac{x}{\lambda}\right) + F\left(\frac{x-\lambda}{\lambda}\right)\right); \tag{4.8}$$

or equivalently, for expansive scaling number $s = 1/\lambda$,

$$F(x) = \frac{1}{2}\left(F(sx) + F(sx-1)\right).$$

(b) Then the following iterative and cascading algorithm holds: Initializing F_0 by setting F_0 to be the Heaviside function

$$F_0(x) = 0 \qquad \text{for } x < 0, \text{ and}$$
$$F_0(x) = 1 \qquad \text{for } x \geq 0,$$

we get the following recursion: F_0, F_1, $\ldots$, etc., with

$$F_{n+1}(x) = \frac{1}{2}\left(F_n(sx) + F_n(sx-1)\right), \qquad n = 0, 1, 2, \ldots, \tag{4.9}$$

or equivalently

$$F_{n+1}(x) = \frac{1}{2}\left(F_n\left(\frac{x}{\lambda}\right) + F_n\left(\frac{x-\lambda}{\lambda}\right)\right).$$

Moreover, $F_n(x) = 1$ holds for $x > \lambda/(1-\lambda)$, and for all n; and the sequence F_n is convergent pointwise in the closed interval

$$X = X_\lambda = \left[0, \frac{\lambda}{1-\lambda}\right] = \left[0, \frac{1}{s-1}\right]. \tag{4.10}$$

For each x, this convergence is monotone, and

$$F(x) = \inf_n F_n(x).$$

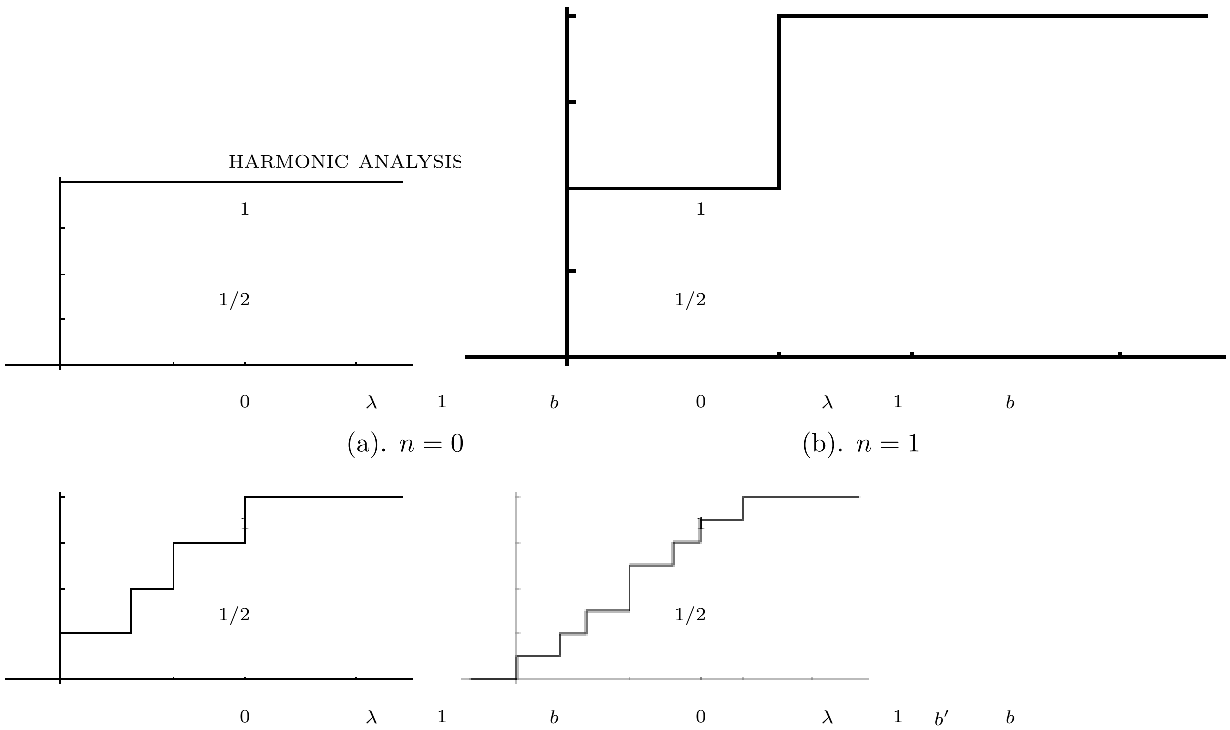
Proof. The scaling identity (4.8) follows from the following facts (see [Hut81]): The limit formula (3.6); the fixed-point property (3.3), i.e., $T(\mu) = \mu$; and the fact that the equilibrium measure μ is supported in the closed interval

$$X = X_\lambda = \left[0, \frac{\lambda}{1-\lambda}\right].$$

See also formula (3.22).

Ad (a): Specifically, for $x \in X_\lambda$, we have

$$\begin{aligned}
F(x) &\underset{\text{by (4.2)}}{=} \int_{X_\lambda} \chi_{[0,x]}(y)\, d\mu(y) \\
&\underset{\text{by (4.1)}}{=} \frac{1}{2}\left(\int_0^{\frac{\lambda}{1-\lambda}} \chi_{[0,x]}(\tau_0(y))\, d\mu(y) + \int_0^{\frac{\lambda}{1-\lambda}} \chi_{[0,x]}(\tau_1(y))\, d\mu(y)\right) \\
&\underset{\text{supp}(\mu)\subset X_\lambda}{=} \frac{1}{2}\left(\int_0^{\frac{x}{\lambda}} d\mu(y) + \int_0^{\frac{x-\lambda}{\lambda}} d\mu(y)\right) \\
&\underset{\text{by (4.2)}}{=} \frac{1}{2}\left(F\left(\frac{x}{\lambda}\right) + F\left(\frac{x-\lambda}{\lambda}\right)\right).
\end{aligned}$$



(a). $n = 0$ (b). $n = 1$

(c). $n = 2$ (d). $n = 3$

FIGURE 1. The series of cascade approximants F_n, $n = 0, 1, 2, \ldots$, to the cumulative distribution function F_λ in the case when $\lambda = \left(\sqrt{5} - 1\right)/2$. The marks on the horizontal axis in addition to λ and 1 are $b' = \lambda + \lambda^2 + \cdots + \lambda^n$, the endpoint of the support of the n'th cascade iteration F_n, and $b = \lambda/(1 - \lambda)$, the endpoint of the support of F_λ. In the case of F_3 for this particular "golden" value of λ, it may be observed that the set of node-points $N_3(\lambda) = \left\{0, \lambda^3, \lambda^2, \lambda, \lambda + \lambda^3, \lambda + \lambda^2, 2\lambda\right\}$, and $\# N_3(\lambda) = 7$, not 8: one of the steps is doubled, due to the fact that $\lambda^2 + \lambda^3 = \lambda$.

For the evaluation of the second integral, note: $\tau_1(y) \in [0, x] \iff -1 \leq y \leq \frac{x-\lambda}{\lambda}$. But since $\operatorname{supp}(\mu) \subset X_\lambda$, we get that

$$\int_{-1}^{\frac{x-\lambda}{\lambda}} d\mu(y) = \int_0^{\frac{x-\lambda}{\lambda}} d\mu(y).$$

Ad (b): Figures 1(a), (b), (c), ... are sketches of the successive functions F_0, F_1, F_2, ... in the case of $\lambda = \frac{\sqrt{5}-1}{2}$, i.e., the reciprocal of the golden ratio $\phi = \left(\sqrt{5} + 1\right)/2$.

As indicated, the sequence is monotone, i.e., $F_{n+1}(x) \leq F_n(x)$ holds; as will be proved.

We now prove the assertion $F_n(x) \equiv 1$ for all n and all $x > \lambda/(1 - \lambda) = \lambda + \lambda^2 + \lambda^3 + \cdots$. As before, $\lambda \in (1/2, 1)$ is fixed, and $\lambda/(1 - \lambda)$ is the right-hand endpoint in the interval X_λ.

The assertion follows by induction. It holds for F_0 since F_0 is the Heaviside function. Now suppose it holds for n. Let $x > \lambda/(1 - \lambda)$ be given. Then

$$\frac{x}{\lambda} > \frac{1}{1 - \lambda} > \frac{\lambda}{1 - \lambda},$$

and

$$\frac{x-\lambda}{\lambda} > \frac{\frac{\lambda}{1-\lambda}-\lambda}{\lambda} = \frac{\lambda}{1-\lambda};$$

so

$$F_{n+1}(x) = \frac{1}{2}\left(F_n\left(\frac{x}{\lambda}\right) + F_n\left(\frac{x-\lambda}{\lambda}\right)\right) = \frac{1}{2}(1+1) = 1,$$

completing the induction.

The separate assertion about pointwise convergence

$$\lim_{n\to\infty} F_n(x) = F(x) \tag{4.11}$$

follows from the stronger fact: For each x, the sequence $(F_n(x))_n$ is monotone, i.e.,

$$F_n(x) \le F_{n-1}(x). \tag{4.12}$$

Induction: Clearly (4.12) holds if $n = 1$. Now suppose it holds up to n. Then by the recursion,

$$F_n(x) - F_{n+1}(x) = \frac{1}{2}\left(\left(F_{n-1}\left(\frac{x}{\lambda}\right) - F_n\left(\frac{x}{\lambda}\right)\right) + \left(F_{n-1}\left(\frac{x-\lambda}{\lambda}\right) - F_n\left(\frac{x-\lambda}{\lambda}\right)\right)\right).$$

The conclusion (4.12) now follows for $n+1$, and the proof is complete. □

4.2. **Measuring overlap.**

Remark 4.7. The three figures Figs. 1(a), (b), (c) taken by themselves offer an oversimplification in that the ordering of the node-points

$$N_n(\lambda) := \left\{ \omega_1\lambda + \omega_2\lambda^2 + \cdots + \omega_n\lambda^n \mid \omega_i \in \{0,1\} \right\}$$

is unique only up to $n = 2$, i.e.,

$$0 < \lambda^2 < \lambda < \lambda + \lambda^2$$

holds for all $\lambda \in (1/2, 1)$. But in general for $n > 2$, the ordering of the 2^n points in $N_n(\lambda)$ is subtle. For example, for $n = 3$, we have

$$\begin{cases} \lambda < \lambda^2 + \lambda^3 & \text{if } \lambda > \dfrac{\sqrt{5}-1}{2}, \\ \lambda = \lambda^2 + \lambda^3 & \text{if } \lambda = \dfrac{\sqrt{5}-1}{2}, \\ \lambda > \lambda^2 + \lambda^3 & \text{if } \lambda < \dfrac{\sqrt{5}-1}{2}, \end{cases} \tag{4.13}$$

indicating that even in N_3 order-reversal may occur depending on the chosen value of λ in $(1/2, 1)$.

The list in (4.13) further shows that the points in each set $N_n(\lambda)$ for $n > 2$ typically occur with multiplicity.

The assertion in (4.13) for the special case $\lambda = \left(\sqrt{5}-1\right)/2$ states that

$$\pi(1000\ldots) = \pi(011000\ldots),$$

where $\pi = \pi_\lambda$ is the encoding mapping $\pi_\lambda \colon \Omega \to X_\lambda$ from Lemma 3.3 and (3.21). While it is known that for all $\lambda \in (1/2, 1)$, π_λ is ∞–1, the infinite sets $\pi_\lambda^{-1}(x)$ are not well understood.

Proposition 4.8. *If* $\lambda = \left(\sqrt{5}-1\right)/2$ *and* $\mu = \mu_\lambda$ *is normalized, i.e.,* $\mu(X) = 1$, *then*

$$\mu\left(\tau_0(X) \cap \tau_1(X)\right) = \frac{1}{3}.$$

Proof of Proposition 4.8. From (4.1) we see that $\tau_0(X) = \left[0, \lambda^2/(1-\lambda)\right]$ and $\tau_1(X) = [\lambda, \lambda/(1-\lambda)]$. A symmetry consideration (Lemma 4.1) further shows that $\mu(\tau_0(X)) = \mu(\tau_1(X))$. We will compute $\mu(\tau_1(X))$ using (3.22). In fact, we show that

$$\mu(\tau_1(X)) = \frac{2}{3}. \tag{4.14}$$

Since $\mu(\tau_0(X) \cup \tau_1(X)) = \mu(X) = 1$, and $1 = \mu(\text{union}) = 2\mu(\tau_1(X)) - \mu(\text{overlap})$, we get

$$\mu\left(\tau_0(X) \cap \tau_1(X)\right) = \frac{4}{3} - 1 = \frac{1}{3},$$

as claimed.

Set $b = b_\lambda = \lambda/(1-\lambda)$. Since $\lambda^2 + \lambda - 1 = 0$, we then get $b = 1/\lambda = 2/\left(\sqrt{5}-1\right)$, and therefore $\tau_1(X) = [\lambda, b]$.

We now turn to (4.14). By (3.22),

$$\mu(\tau_1(X)) = \lim_{n\to\infty} \left(2^{-n}\right) \cdot \#\left(\pi^{-1}\left(N_n(\lambda) \cap [\lambda, b]\right)\right).$$

Recall that

$$N_n(\lambda) = \left\{ \sum_{i=1}^{n} \omega_i \lambda^i \;\middle|\; \omega = (\omega_1 \dots \omega_n) \in \{0,1\}^n \right\}$$

for all $n \in \mathbb{N}$. We claim that for $n > 2$,

$$\begin{aligned} \pi^{-1}\left(N_n(\lambda) \cap [\lambda, b]\right) = \{\, \omega \in \{0,1\}^n \mid \omega_1 = 1 \,\} \\ \cup \{\, \omega \in \{0,1\}^n \mid \omega_1 = 0,\ \omega_2 = \omega_3 = 1 \,\} \cup \cdots, \end{aligned} \tag{4.15}$$

where the union of the individual sets on the right-hand side is clearly disjoint. The sets indicated by "$\cup \cdots$" on the right-hand side in (4.15) have the form

$$\bigcup_k \{\, \omega \mid \omega_{2i-1} = 0,\ \omega_{2i} = 1,\ i \leq k;\ \text{and}\ \omega_{2k+1} = 1 \}.$$

The contribution to these sets is $\geq \lambda^2 + \lambda^4 + \lambda^6 + \cdots + \lambda^{2k} + \lambda^{2k+1} = \lambda$, where we used that $\lambda + \lambda^2 = 1$. It follows that

$$\lim_{n\to\infty} \left(2^{-n}\right) \cdot \#\left(\pi^{-1}\left(N_n(\lambda) \cap [\lambda, b]\right)\right) = 2^{-1} + 2^{-3} + 2^{-5} + 2^{-7} + \cdots = \frac{2}{3}$$

as claimed.

It is clear that the right-hand side in (4.15) is contained in $\pi^{-1}\left(N_n(\lambda) \cap [\lambda, b]\right)$ for sufficiently large $n > 2$. The assertion that they are equal follows from the fact that

$$\lambda^3 + \lambda^4 + \cdots + \lambda^m < \lambda \tag{4.16}$$

for all $m > 3$. But using $1 - \lambda = \lambda^2$, note that (4.16) is equivalent to $1 - \lambda^{m-2} < 1$, which is clearly true. This proves the proposition.

Recall that the overlap in μ_λ-measure is

$$\mu_\lambda\left([\lambda, b(\lambda) - \lambda]\right) = P_{1/2}\left(\{\, \omega \in \Omega \mid \pi_\lambda(\omega) \in [\lambda, b(\lambda) - \lambda] \,\}\right).$$

This means that the contributions to $\pi_\lambda^{-1}([\lambda, b(\lambda)-\lambda])$ with $P_{1/2}$-measure equal to zero may be omitted in the accounting (4.15) above.

However, even for $\lambda = \left(\sqrt{5}-1\right)/2$, even the infinite set $\pi_\lambda^{-1}(\{\lambda\})$ has interesting dynamics. But its contribution to the overlap in μ-measure is

$$\mu_\lambda(\{\lambda\}) = 0;$$

see also Remark 4.4.

Notation. Set

$$\begin{aligned} w &:= (0\,1), \\ \underline{0} &:= (\underbrace{0\,0\,0\,\dots}_{\infty\text{ repetition}}), \text{ and} \\ \underline{1} &:= (\underbrace{1\,1\,1\,\dots}_{\infty\text{ repetition}}). \end{aligned}$$

Then $\pi_\lambda^{-1}(\{\lambda\})$ contains the following infinite lists:

$$\begin{aligned} &(1\,\underline{0}), \\ &(w\,1\,\underline{0}), \\ &(w\,w\,1\,\underline{0}), \\ &(w\,w\,w\,1\,\underline{0}), \\ &\vdots, \end{aligned}$$

etc., and

$$\begin{aligned} &(0\,0\,\underline{1}), \\ &(w\,0\,0\,\underline{1}), \\ &(w\,w\,0\,0\,\underline{1}), \\ &(w\,w\,w\,0\,0\,\underline{1}), \\ &\vdots, \end{aligned}$$

etc. □

Remark 4.9. The fact that for $\lambda = \left(\sqrt{5}-1\right)/2$ ($= \phi^{-1}$, ϕ = golden ratio) the overlap measured in the Hutchinson measure is $1/3$ appears also to follow from [SV98]. In [SV98, Cor. 1.2, p. 220], entirely about the golden shift, Sidorov and Vershik do get $1/3$ by a completely different argument: they introduce a transition matrix on a combinatorial tree, and when translated into our $\lambda = \left(\sqrt{5}-1\right)/2$ example, their Sidorov–Vershik tree is then Fibonacci. That is key to their computations.

In contrast, our method is general and applies to general metric spaces: IFSs with overlap. Even when specialized to 1D, for the special case of (4.1), our method has the advantage (see Corollary 4.11 and Remark 4.12 below) of estimating the overlap in Hutchinson measure also when λ is not ϕ^{-1}, i.e., is not "golden."

Remarks 4.10. (a) The Lebesgue measure of the intersection $\tau_0(X) \cap \tau_1(X)$ is $\frac{\lambda^2}{1-\lambda} - \lambda$, which for $\lambda = \frac{\sqrt{5}-1}{2}$ works out to

$$\operatorname{Leb}\left(\tau_0(X) \cap \tau_1(X)\right) = \frac{3-\sqrt{5}}{2}.$$

(b) Since

$$\frac{3-\sqrt{5}}{2} > \frac{1}{3},$$

we conclude from the proposition that the Hutchinson measure of the intersection is the smaller of the two.

(c) The argument from the proof of the proposition extends to the IFS $\tau_k^{(\lambda)}(x) := \lambda(x+k)$, $k \in \{0,1\}$, for all $\lambda \in (1/2, 1)$, and we conclude that there is essential overlap for all λ in $(1/2,1)$, but an explicit formula for $\mu_\lambda\left(\tau_0^{(\lambda)}(X_\lambda) \cap \tau_1^{(\lambda)}(X_\lambda)\right)$ (> 0) is not known.

The following is a consequence of the argument in the proof of Proposition 4.8.

Corollary 4.11. (a) *For all* $\lambda \in \left[\left(\sqrt{5}-1\right)/2, 1\right)$ *we have*

$$\mu_\lambda\left(\tau_0^{(\lambda)}(X_\lambda) \cap \tau_1^{(\lambda)}(X_\lambda)\right) \geq \frac{1}{3}. \tag{4.17}$$

(b) *For all* $\lambda \in \left(1/2, \left(\sqrt{5}-1\right)/2\right)$, *there is some* m, *depending on* λ, *such that*

$$\lambda + \lambda^2 + \lambda^3 + \cdots + \lambda^m \geq 1, \tag{4.18}$$

and for such a choice of m *we have*

$$\mu_\lambda\left(\tau_0^{(\lambda)}(X_\lambda) \cap \tau_1^{(\lambda)}(X_\lambda)\right) \geq \frac{1}{2^m - 1}. \tag{4.19}$$

Remark 4.12. If $\lambda = \left(\sqrt{5}-1\right)/2$, the number m in (4.18) may be taken to be $m = 2$, and in that case (4.18) is "=".

Proof of Corollary 4.11. *Ad* (a): An easy modification of the argument in the proof of Proposition 4.8 shows that if $\lambda^2 + \lambda > 1$, then (4.17) holds.

Ad (b): Let $\lambda \in \left(1/2, \left(\sqrt{5}-1\right)/2\right)$ be given. It follows from algebra that if m is sufficiently large, then (4.18) must hold. We pick m to be the first number which gets the sum on the left-hand side in (4.18) ≥ 1.

Consider the following specific finite word w in $\{0,1\}^{\text{finite}}$ given by

$$w = (\,0\,\underbrace{1\,1\,\ldots\,1}_{m-1\text{ times}}\,)$$

and generate more words as follows:

$$\begin{aligned}
&(\,1\,\text{free}\,),\\
&(\,w\,1\,\text{free}\,),\\
&(\,w\,w\,1\,\text{free}\,),\\
&(\,w\,w\,w\,1\,\text{free}\,),\\
&\text{etc.},
\end{aligned}$$

where "free" means unrestricted strings of bits. As before, the resulting sequence of subsets in Ω is disjoint. Then

$$\mu_\lambda\left(\tau_1^{(\lambda)}\left(X_\lambda\right)\right) \geq 2^{-1}+2^{-m-1}+2^{-2m-1}+2^{-3m-1}+\cdots = 2^{-1}\frac{1}{1-2^{-m}} = \frac{2^{m-1}}{2^m-1}.$$

The argument from the proposition yields

$$\mu_\lambda\left(\tau_0^{(\lambda)}\left(X_\lambda\right)\cap\tau_1^{(\lambda)}\left(X_\lambda\right)\right) \geq \frac{2^m}{2^m-1}-1 = \frac{1}{2^m-1},$$

which is the assertion (4.19). □

Cascade approximation. Each function $F_n^{(\lambda)}$ from the approximation to the cumulative distribution in Proposition 4.6 has a finite set of node points $N_n(\lambda)$, and

$$\# N_n(\lambda) \leq 2^n \qquad \text{for all } n;$$

but for fixed λ, the configuration of points in $N_n(\lambda)$ can be complicated for $n>2$ and large; and each set $N_n(\lambda)$ also depends on the particular numerical choice of a value for λ.

This is borne out in the cascade of figures (see Figure 1) made for $\lambda=(\sqrt{5}-1)/2$. We have included pictures of $F_1^{(\lambda)}$, $F_2^{(\lambda)}$, $\dots$, up to $F_4^{(\lambda)}$. As sketched in Remark 4.7, the reason is that the detailed configuration and the multiplicities in the sets $N_n(\lambda)$ of node points are reflected in the progression of graphs of the functions $F_n^{(\lambda)}(\,\cdot\,)$. This fine structure only becomes visible for large $n>2$.

In Figure 2 above, we summarize the distribution of $\pi_\lambda(\,\cdot\,)$, i.e., the function $F_\lambda(\,\cdot\,)$ in (4.2). By [Sol95], $F_\lambda(\,\cdot\,)$ is only known to have L^1-a.e. derivative, or L^1-density, for a.e. λ in $1/2<\lambda<1$.

The differences in cascade approximation when the value of λ varies is illustrated by Figures 1, 2, 3, and 4. We have already outlined that $\lambda=(\sqrt{5}-1)/2$ is special. Figure 1 illustrates the cascades from Proposition 4.6, $F_0^{(\lambda)} \to F_1^{(\lambda)} \to \dots \to F_7^{(\lambda)}$, with each step representing a λ scaled subdivision. Figure 2 represents the "idealized" limit of the iteration. Since there are only very few rigorous results in the literature for specific values of λ in the open interval $(1/2,1)$ (see Section 1), we have included Mathematica sketches of of F_λ for two chosen values of λ in Figures 3 and 4.

Our next result gives the two Radon–Nikodym derivatives in the case of (4.1) for λ fixed, i.e., $1/2<\lambda<1$ is chosen. The measure $\mu=\mu_\lambda$ is chosen such that $\mu_\lambda(X_\lambda)=1$. Then set

$$\varphi_i := \frac{d\mu\circ\tau_i^{-1}}{d\mu}, \qquad i=0,1. \tag{4.20}$$

Proposition 4.13. *The two Radon–Nikodym derivatives φ_0 and φ_1 from* (4.20) *are given by the formulas*

$$\frac{1}{2}\,\varphi_0(x) = \begin{cases} 1 & \text{if } 0\leq x<\lambda, \\ F\left(\dfrac{\lambda^2-x(1-\lambda)}{(1-\lambda)(2\lambda-1)}\right) & \text{if } \lambda\leq x\leq\dfrac{\lambda^2}{1-\lambda}, \\ 0 & \text{if } \dfrac{\lambda^2}{1-\lambda}<x\leq\dfrac{\lambda}{1-\lambda}, \end{cases} \tag{4.21}$$

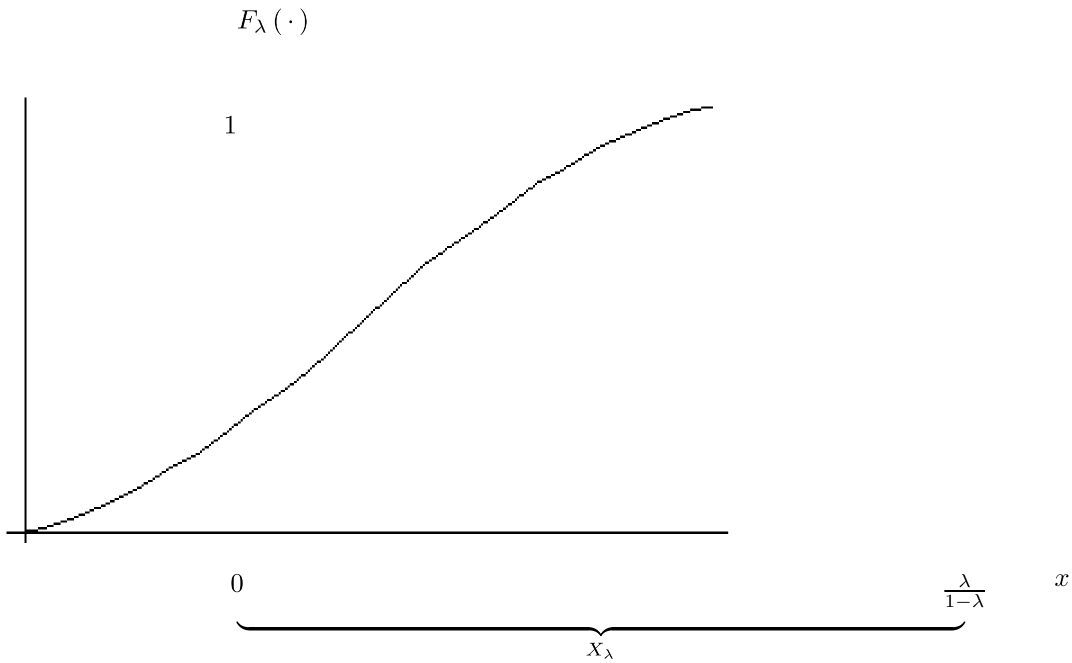


FIGURE 2. The cumulative distribution of $\pi_\lambda\left(\,\cdot\,\right)$, where $\lambda = (\sqrt{5}-1)/2$. Caution: A closed formula for $F_\lambda\left(\,\cdot\,\right)$ is not known. But using the second formula in (4.1) and (4.2), the reader may check that for fixed λ, the cumulative distribution F $(= F_\lambda)$ satisfies the scaling identity $F(x) = \frac{1}{2}(F(x/\lambda) + F((x-\lambda)/\lambda))$.

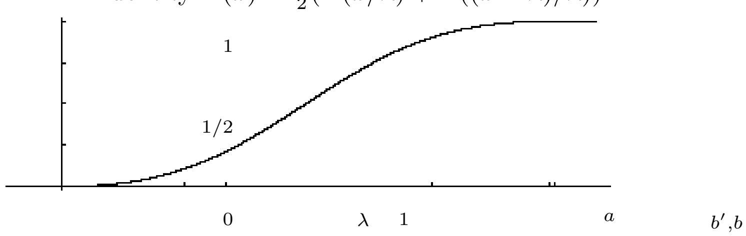


FIGURE 3. $\lambda = 3/4$

and

$$\frac{1}{2}\,\varphi_1\left(x\right) = \begin{cases} 0 & \text{if } 0 \leq x < \lambda, \\ F\left(\dfrac{(x-\lambda)}{(2\lambda-1)}\right) & \text{if } \lambda \leq x \leq \dfrac{\lambda^2}{1-\lambda}, \\ 1 & \text{if } \dfrac{\lambda^2}{1-\lambda} < x \leq \dfrac{\lambda}{1-\lambda}. \end{cases} \tag{4.22}$$

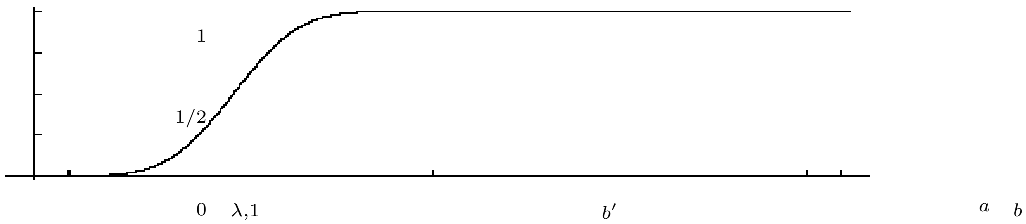


FIGURE 4. $\lambda = 23/24$

Proof. The result follows from Corollary 3.5 and from the present discussion of Example 3.7. $\square$

Remark 4.14. For the convenience of the reader, we have sketched the two functions φ_0 and φ_1 in Figures 5 and 6 above.

Details for (4.22), *Figure* 6. Using (3.22), (4.2), Corollary 3.5, and conditional probabilities, we get for all $S \in \mathcal{B}$ (Borel subsets of X_λ) the formula

$$\mu\left(\tau_1^{-1}(S)\right) = \int_S P\left(\{\pi_\lambda \circ \sigma_1 \leq x\}\right) d\mu(x), \tag{4.23}$$

where the function under the integral sign in (4.23) coincides with the expression in (4.22). $\square$

Remark 4.15. As above, let $1/2 < \lambda < 1$, and $X_\lambda = \left[0, \lambda/(1-\lambda)\right]$. Then $\tau_0(X_\lambda) = \left[0, \lambda^2/(1-\lambda)\right]$ and $\tau_1(X_\lambda) = \left[\lambda, \lambda/(1-\lambda)\right]$; see Figures 5 and 6. We include this note to stress that the Radon–Nikodym derivatives are sensitive to the choice of μ. Take for example $\mu =$ Lebesgue measure on X_λ. Then an easy computation yields

$$\frac{d\mu \circ \tau_i^{-1}}{d\mu} = \frac{1}{\lambda}\chi_{\tau_i(X_\lambda)}, \qquad i = 0, 1;$$

but the corresponding $\mathbb{F}$ is then *not* a column isometry.

Proposition 4.16. *Continuing the example in Proposition* 4.13*. If we choose* μ *to be the equilibrium measure, then the isometry* $\mathbb{F} = \begin{pmatrix} F_0 \\ F_1 \end{pmatrix} : L^2(\mu) \to L^2(\mu)_2$ *yields the following formula for the range-projection in* $L^2(\mu)_2$*:*

$$\mathbb{F}\mathbb{F}^* = \begin{pmatrix} F_0F_0^* & F_0F_1^* \\ F_1F_0^* & F_1F_1^* \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \varphi_0 \circ \tau_0 & (\varphi_1 \circ \tau_0)\, T_{-1} \\ (\varphi_0 \circ \tau_1)\, T_1 & \varphi_1 \circ \tau_1 \end{pmatrix}, \tag{4.24}$$

where the composite functions $\varphi_i \circ \tau_j$ *serve as multiplication operators in* $L^2(\mu)$*, while the two other operators making up the entries in* (4.24) *are*

$$(T_{\pm 1} f)(x) = f(x \pm 1), \qquad f \in L^2(\mu).$$

As before λ *is fixed such that* $1/2 < \lambda < 1$*, and* $\mu = \mu_\lambda$ *is the equilibrium measure.*

Proof. The result follows from a computation, and an application of Remark 2.6 and Lemma 2.8. $\square$

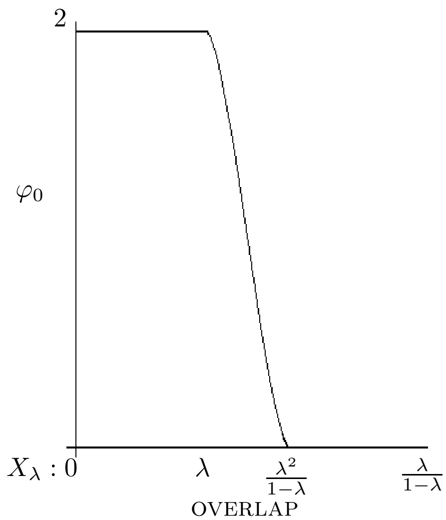


FIGURE 5. The Radon–Nikodym derivative φ_0.

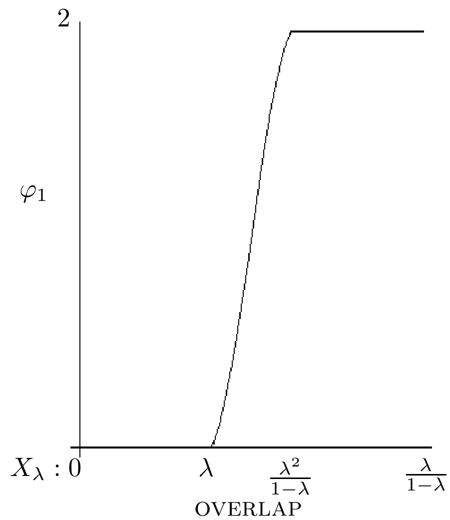


FIGURE 6. The Radon–Nikodym derivative φ_1.

Corollary 4.17. *Continuing the example from Propositions* 4.13 *and* 4.16, *we note that the function*

$$\mathbf{f} = \begin{pmatrix} f_0(x) \\ f_1(x) \end{pmatrix} = \begin{pmatrix} \varphi_1 \circ \tau_0 \\ -\varphi_0 \circ \tau_1 \end{pmatrix} \in L^2(\mu)_2 \tag{4.25}$$

is nonzero, and it is in the orthogonal complement of the range of the column isometry $\mathbb{F}$.

Proof. Using (4.24), a computation shows that $\mathbf{f}$ in (4.25) satisfies $\mathbb{F}\mathbb{F}^*\mathbf{f} = 0$; and moreover that $\mathbf{f} = \begin{pmatrix} f_0 \\ f_1 \end{pmatrix}$ is nonzero in $L^2(\mu)_2$. □

5. ESSENTIAL OVERLAP AND GAPS FOR SIERPINSKI CONSTRUCTIONS IN $\mathbb{R}^2$

This section includes a 2D variant of the 1D examples from Section 4 above. The case of 2D is interesting and different in that the fractal features become apparent both for the associated Hutchinson measures, as well as for their support. Contrast: In the best known gap fractal, the Cantor set, the middle thirds are omitted, and as a result the familiar cascading gap-subintervals emerge. As we saw in Example 3.7, in 1D (CASE 3) when λ is adjusted, $\lambda > 1/2$ so as to create μ_λ-essential overlap for the infinite convolution IFS, then the omitted middles disappear, and the resulting X_λ is simply an interval. Not so for the analogous 2D construction! As we see, in 2D overlap and gaps may coexist!

Recall that the attractor for a contractive IFS in a complete metric space is equal to the support of the corresponding Hutchinson measure.

In the next result, we show that when the same procedure from Proposition 4.6 (and (4.1)–(4.2)) is extended to 2D we get a one-parameter family of fractals with gaps, and essential overlap at the same time. So in the context of (3.2) and (3.3), the ambient space $Y = \mathbb{R}^2$, the weights in (3.3) are $p_i = 1/3$, the number λ will

(a). $n = 1$

(b). $n = 2$

FIGURE 7. The first two iterations of X_λ for $\lambda = 11/20$.

be in the open interval $(1/2, 2/3)$, and the attractor $X = X_\lambda$ will be a Sierpinski fractal with essential overlap and gaps. Its Hausdorff dimension H_λ will be

$$H_\lambda = -\frac{\log 3}{\log \lambda}.$$

This Sierpinski fractal $= X_\lambda$ is sketched below in Figures 7, 8, 9, and 10 for the cases $\lambda = 11/20$, $\lambda = \left(\sqrt{5}-1\right)/2$, $\lambda = 13/20$, and $\lambda = 3/4$, respectively.

Let λ be fixed in the open interval $(1/2, 2/3)$. The case $\lambda = 1/2$ is the familiar Sierpinski gasket in 2D.

Recursion: By analogy to the middle-third Cantor-set construction, start with a triangle T_λ with vertices $(0,0)$, $(\lambda/(1-\lambda))\, u_1$, and $(\lambda/(1-\lambda))\, u_2$.

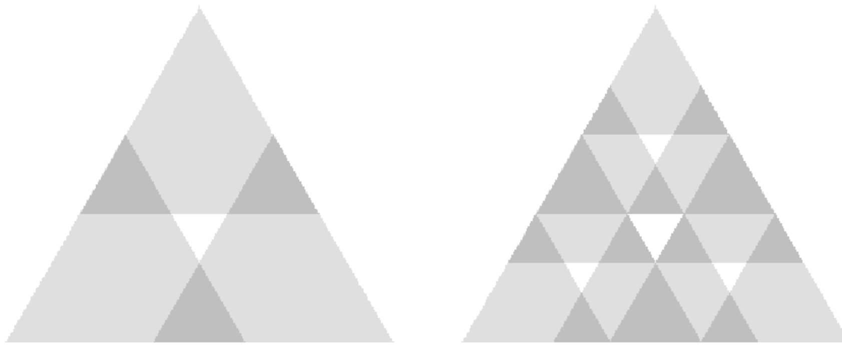
The three λ-scaled triangles $\tau_i(T_\lambda)$ are shaded in light grayscale, and they have pairwise overlaps as indicated in the first iteration in Fig. 8. These three pairwise overlaps are the first smaller dark-shaded triangles. The recursion now continues.

Omitted triangles: The first gap, i.e., the first white interior triangle, is the set-theoretic difference $G_1 = T_\lambda \backslash \bigcup_i \tau_i(T_\lambda)$. The fractal X_λ now arises by iteration just as in the familiar case of recursion for the middle-third Cantor set. In the first step of the recursion, there is just one interior and centered gap-triangle; it is inverted from the position of the initial T_λ.

Also note that the size of the overlaps in the iteration is monotone in λ, small when λ is close to $1/2$, the usual Sierpinski gasket. In contrast, the omitted middles decrease gradually with λ and they collapse to points when $\lambda = 2/3$. The nature of the overlaps changes at the value $\lambda = (\sqrt{5}-1)/2$. We will discuss the nature of the overlaps in Section 5.2.

5.1. **The 2D recursion.** Initialize the recursion sequence at level $n = 0$ with an inflated triangle T_λ, inflation factor $= \lambda/(\lambda-1)$, and then generate the fractal X_λ sequentially, $n = 1, 2, \ldots$, by the usual iteration limit,

$$X_\lambda = \overline{\bigcap_{n=1}^{\infty} \bigcup_{w:\ \text{word of length } n} \tau_w(T_\lambda)}.$$



(a). $n = 1$ (b). $n = 2$

FIGURE 8. The first two iterations of X_λ for $\lambda = (\sqrt{5} - 1)/2$. See Example 5.1.

Here τ_w denotes an n-fold composition of the individual τ_i maps with the indices I making up the word w, i.e., w giving the address of the respective "small" triangles $\tau_w(T_\lambda)$ for the n'th level iteration:

$$\tau_i(x) = \lambda(x + u_i), \qquad i = 0, 1, 2,\ x \in \mathbb{R}^2,$$
$$\tau_w = \tau_{i_1} \cdots \tau_{i_n} \qquad \text{if } w = (i_1, \dots, i_n).$$

Example 5.1. 2D Sierpinski with overlap and gaps, $1/2 < \lambda < 2/3$**.** Let the vectors u_0, u_1, u_2 in $\mathbb{R}^2$ be given as follows:

$$\begin{cases} u_0 = (0,0), \\ u_1 = (1,0), \text{ and} \\ u_2 = \left(\dfrac{1}{2}, \dfrac{\sqrt{3}}{2}\right), \text{ and set} \end{cases} \tag{5.1}$$

$$\Omega = \{u_0, u_1, u_2\}^{\mathbb{N}} \tag{5.2}$$

with Bernoulli measure $P_{1/3}$ corresponding to the infinite-product measure $p_i = 1/3$, $i = 0, 1, 2$, as in (3.3) and Corollary 3.4. We will denote by μ_λ the measure in Corollary 3.5, and we set $X_\lambda := \operatorname{supp}(\mu_\lambda)$.

The IFS is

$$\begin{cases} \tau_0^{(\lambda)}(x) = \lambda x, \\ \tau_1^{(\lambda)}(x) = \lambda(x + u_1), \text{ and} \\ \tau_2^{(\lambda)}(x) = \lambda(x + u_2) \end{cases} \tag{5.3}$$

for $x \in \mathbb{R}^2$. Here we make the restriction $1/2 < \lambda < 2/3$. As noted in Section 3,

$$X_\lambda = \bigcup_{i=0}^{2} \tau_i^{(\lambda)}(X_\lambda) \tag{5.4}$$

and X_λ is the unique compact ($\neq \varnothing$) solution to (5.4).

Let $\Omega := \{0, u_1, u_2\}^{\mathbb{N}}$, and let $P_{1/3}$ be the usual Bernoulli measure on Ω with equal and independent probabilities $(1/3, 1/3, 1/3)$. The formula (3.21) for the random variable $\pi_\lambda : \Omega \to X_\lambda$ extends from 1D to 2D with the only modification that the coefficients ω_i from (3.21) now take values in the finite alphabet of vectors $\{0, u_1, u_2\}$.

Let A_i^λ, $i = 0, 1, 2$ be the three vertices in X_λ (Figure 8); i.e. $A_0^\lambda = (0, 0)$, $A_1^\lambda = \frac{\lambda}{1-\lambda} u_1$, and $A_2^\lambda = \frac{\lambda}{1-\lambda} u_2$. Note that for each $\lambda \in (1/2, 1)$, X_λ is contained in the triangle T_λ with the vertices A_i^λ, $i = 0, 1, 2$.

Our first result concerns symmetry, and it is an immediate extension of Lemma 4.1 from the 1D case to the 2D case. For each $i \in \{0, 1, 2\}$, let $S_i^\lambda(x)$ denote the equilateral triangle of side-length x with vertex A_i^λ, which shares two sides with segments of sides in T_λ. Then the argument from Lemma 4.1 shows that for each $x \in \mathbb{R}_+$, the three numbers

$$P_{1/3}\{\omega \in \Omega | \pi_\lambda(\omega) \in S_i^\lambda(x)\}$$

coincide. Since $\mu_\lambda = P_{1/3} \circ \pi_\lambda^{-1}$ by Corollary 3.5, we conclude in particular that the three numbers $\mu_\lambda(\tau_i^{(\lambda)}(X_\lambda)$ agree for $i = 0, 1, 2$.

For $i \neq j$, set

$$OV_{ij}^\lambda := \tau_i^{(\lambda)}(X_\lambda) \cap \tau_j^{(\lambda)}(X_\lambda).$$

It follows that

$$\mu_\lambda(OV_{01}^\lambda) = \mu_\lambda(OV_{02}^\lambda) = \mu_\lambda(OV_{12}^\lambda).$$

Proposition 5.2. (a) *For $\lambda \in (1/2, 2/3)$, the fractal X_λ has Hausdorff dimension*

$$H_\lambda = -\frac{\log 3}{\log \lambda}. \tag{5.5}$$

It has essential overlap

$$\mu_\lambda \left(\tau_i^{(\lambda)}(X_\lambda) \cap \tau_j^{(\lambda)}(X_\lambda) \right) > 0 \qquad \text{for } i \neq j \tag{5.6}$$

and the μ_λ-measure of the pairwise overlaps is independent of the pair (i, j) with $i \neq j$.

(b) *Setting*

$$\pi_\lambda(\omega) = \sum_{i=1}^{\infty} \omega_i \lambda^i, \tag{5.7}$$

$\omega = (\omega_i)_1^\infty \in \Omega$, we get a vector-valued random variable, and

$$\mu_\lambda = P_{1/3} \circ \pi_\lambda^{-1}$$

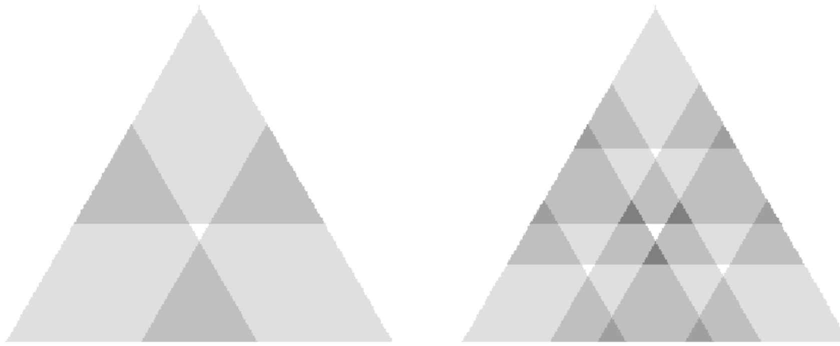
holds, where $\pi_\lambda \colon \Omega \to X_\lambda$ is the encoding mapping.

(c)

$$\tau_0^{(\lambda)}(X_\lambda) \cap \tau_1^{(\lambda)}(X_\lambda) \cap \tau_2^{(\lambda)}(X_\lambda) = \varnothing.$$

Proof. Since the proof is essentially contained in the previous sections, we shall be brief.

Ad (a): The formula (5.5) in (a) for the Hausdorff dimension follows from the arguments in [CI06].



(a). $n = 1$ (b). $n = 2$

FIGURE 9. The first two iterations of X_λ for $\lambda = \frac{13}{20}$. Note that $\frac{2}{3} - \frac{13}{20} = \frac{1}{60}$, and the gaps are very small.

We begin with the case $\lambda = (\sqrt{5} - 1)/2$, and for the moment, we will drop λ from the notation. Our claim is that

$$\mu(OV_{01}) = \mu(OV_{02}) = \mu(OV_{12}) = \frac{1}{24}.$$

To see this, we introduce the Bernoulli space Ω and consider the cylinder sets in Ω indexed by finite words in the alphabet $\{u_0, u_1, u_2\}$, $u_0 = 0$. If $w = (u_{i_1}\, u_{i_2} \ldots u_{i_n})$, set

$$\Omega(w) := \{\omega \in \Omega | \omega_1 = u_{i_1}, \ldots, \omega_n = u_{i_n}\},$$

and note that $P_{1/3}(\Omega(w)) = 3^{-n}$.

Set $w = (u_0 u_1) = (0\, u_1)$, and consider the following sequence of disjoint cylinder sets:

$$\Omega(u_1), \Omega(w\, u_1), \Omega(w\, w\, u_1), \ldots, \Omega(\underbrace{w\, w \cdots w}_{k \text{ times}}\, u_1), \ldots .$$

The argument from Section 4 shows that

$$\mu(\tau_1(X)) = \sum_{k=0}^{\infty} P_{1/3}\Big(\Omega(\underbrace{w\, w \cdots w}_{k \text{ times}}\, u_1)\Big) = \sum_{k=0}^{\infty} \Big(\frac{1}{3}\Big)^{2k+1} = \frac{3}{8}.$$

Using the symmetry argument from above, we also have

$$1 = \mu\Big(\cup_i \tau_i(X)\Big) = 3\mu(\tau_1(X)) - 3\mu(OV_{01}) = \frac{9}{8} - 3\mu(OV_{01}).$$

The desired result $\mu(OV_{01}) = 1/24$ follows.

Starting at $\lambda_1 = (\sqrt{5}-1)/2$ we see that the function $\lambda \mapsto \mu_\lambda(OV_{01}^\lambda)$ is increasing. Hence, to prove (5.6), we need only establish a lower bound for values of λ in the open interval $(1/2, (\sqrt{5} - 1)/2)$. But this can be done *mutatis mutandis* as in the proof of Corollary 4.11; see (4.19). If $\lambda \in (1/2, (\sqrt{5}-1)/2)$, determine $m \in \mathbb{N}$ as in (4.16). Then it follows that

$$3\mu_\lambda(OV_{01}^\lambda) \geq \frac{1}{3^m - 1}.$$

Ad (b): The proof of (5.6) in (b) is based on symmetry considerations (Lemma 4.1) extended from 1D to 2D, as well as the estimates in Proposition 4.8, Remarks 4.10, and Corollary 4.11.

Ad (c): To see that the triple overlap is empty, calculate the distances between pairwise overlaps to the third sub-partition. □

We conclude this section with the following open question: for what values of λ in the interval $(2/3, 1)$ is μ_λ absolutely continuous with respect to the 2-dimensional Lebesgue measure? There are two pieces of partial evidence for absolute continuity of the two-dimensional μ_λ for a.e. $\lambda \in (2/3, 1)$:

(1) Since the interior gaps close at $\lambda = 2/3$, so that for $\lambda \geq 2/3$, $X_\lambda = T_\lambda$ (the closed triangle), following the 1D analogy, it seems reasonable to expect the a.e. conclusion in this range of λ.
(2) One of the proofs [PS96] in the literature for the 1D case introduces a clever Fubini-Tonelli argument with a function in several variables with λ as one of the integration variables. Using a density argument, one then gets finiteness of a corresponding key functions for a.e. λ in the interval, and this in turn (following the 1D analogy) is likely to yield an expression in 2D for the Radon-Nikodym derivative for those values of λ.

    In particular, for $x \in T_\lambda$, let $L_\lambda^\ell(x)$ be the equilateral triangle centered at x with side length ℓ and area $\frac{\sqrt{3}}{4}\ell^2$. Define

$$\underline{D}(\mu_\lambda, x) := \liminf_{\ell \downarrow 0} \frac{\mu_\lambda(L_\lambda^\ell(x))}{\frac{\sqrt{3}}{4}\ell^2}$$

    (an analogue of the first formula on [PS96, p. 233]). The argument from [PS96] is likely to yield

$$\int_{2/3}^{1} \iint_{T_\lambda} \underline{D}(\mu_\lambda, x)\, d\mu_\lambda(x)\, d\lambda < \infty.$$

    From that we could conclude that $\lambda \mapsto \underline{D}(\mu_\lambda, x)$ is finite for a.e. $\lambda \in (2/3, 1)$, putting the Radon-Nikodym derivative of μ_λ with respect to two-dimensional Lebesgue measure in L^2 for a.e. λ.

5.2. **The nature of the overlaps and induced systems.** As we noted before, the nature of the overlaps changes at the value $\lambda = (\sqrt{5} - 1)/2$. Here we refer to overlaps of monomials in the τ_i's of degree $n = 1, 2, \ldots$ applied to the initial triangle $T = T_\lambda$ as $\tau^n(T)$. Let $\mathbf{ov}(\tau^n(T))$ denote overlaps at level n— for example,

$$\mathbf{ov}(\tau^1(T)) = \Big(\tau_0(T) \cap \tau_1(T)\Big) \cup \Big(\tau_0(T) \cap \tau_2(T)\Big) \cup \Big(\tau_1(T) \cap \tau_2(T)\Big).$$

(i) **The Sierpinski Gasket.** When $\lambda = 1/2$, the resulting fractal X_λ is the Sierpinski gasket, and the essential overlap is a set of Lebesgue measure zero [Str06].
(ii) **Simple Overlap.** When $\lambda \in (1/2, (\sqrt{5} - 1)/2)$, we have

$$\mathbf{ov}(\tau^n(T)) \cap \mathbf{ov}(\tau^{n+1}(T)) = \varnothing.$$

    We call this type of overlap "overlap of multiplicity one" or "simple overlap." When simple overlap occurs, the subset of X_λ which consists of the overlaps

is itself an IFS. So, an IFS with essential but simple overlap induces a new IFS with non-essential overlap. See Figure 7.

(iii) **The Golden 2D Fractal.** Let $\lambda = (\sqrt{5}-1)/2$, as in Figure 8. We see $\mathbf{ov}(\tau^1(T))$ as the dark shaded triangles in Figure 8, picture (a). As we move from (a) to (b) in Figure 8, we see that successive overlap sets intersect at vertices: $\mathbf{ov}(\tau^1(T)) \cap \mathbf{ov}(\tau^2(T))$ is non-empty. In fact, $\mathbf{ov}(\tau^1(T))$ consists of three triangles, and $\mathbf{ov}(\tau^2(T))$ of 9. Each triangle in $\mathbf{ov}(\tau^1(T))$ shares a vertex with one or two from $\mathbf{ov}(\tau^2(T))$, and the double-sharing happens for the interior triangles from $\mathbf{ov}(\tau^2(T))$. This pattern continues, so that at each step of the iteration, each vertex of a triangle in $\mathbf{ov}(\tau^n(T))$ coincides with some vertex in a triangle from $\mathbf{ov}(\tau^{n+1}(T))$. We therefore have an induced system which forms a graph with edges and vertices, with each vertex from $\mathbf{ov}(\tau^n(T))$ connecting to one or two vertices from $\mathbf{ov}(\tau^{n+1}(T))$.

In the following when we refer to disjoint pairs of triangles, we will mean "disjointness of the respective interiors," thus allowing the sharing of vertices. The triangles in $\mathbf{ov}(\tau^n(T))$ are disjoint from all the triangles in $\mathbf{ov}(\tau^{n+1}(T))$, but triangles from $\mathbf{ov}(\tau^{n+2}(T))$ may be contained in triangles from $\mathbf{ov}(\tau^n(T))$. In fact, a triangle from $\mathbf{ov}(\tau^{n+2}(T))$ is either disjoint from triangles in the set $\mathbf{ov}(\tau^n(T))$, or that triangle is contained in a unique triangle from $\mathbf{ov}(\tau^n(T))$.

We can formalize the overlap between iterations by defining a set operation OV which takes a set S to the set

$$\mathrm{OV}(S) := \Big(\tau_0(S) \cap \tau_1(S)\Big) \cup \Big(\tau_0(S) \cap \tau_2(S)\Big) \cup \Big(\tau_1(S) \cap \tau_2(S)\Big).$$

We have already seen that

$$\mathrm{OV}(T_\lambda) = \mathbf{ov}(\tau^1(T)),$$

and we also have

$$\mathrm{OV}(\mathbf{ov}(\tau^1(T))) = \mathbf{ov}(\tau^2(T)).$$

Suppose $\xi = (i_1, i_2, \dots i_n) \in \{0,1,2\}^n$—that is, ξ is a multi-index of length n each of whose components is 0, 1, or 2. We can use ξ to keep track of the monomials in the τ_i's which we mentioned above:

$$\tau_\xi(x) := \tau_{i_1}\tau_{i_2}\cdots\tau_{i_n}(x) = \lambda^n x + \lambda^n u_{i_1} + \lambda^{n-1}u_{i_2} + \dots \lambda u_{i_n}.$$

Then

$$\mathbf{ov}(\tau^{n+1}(T)) = \Big\{\mathrm{OV}(\tau_\xi(T)) : \xi \in \{0,1,2\}^n\Big\}.$$

(iv) **The Residual Interval.** When $(\sqrt{5}-1)/2 < \lambda < 2/3$, the set $\mathbf{ov}(\tau^n(T)) \cap \mathbf{ov}(\tau^{n+1}(T))$ no longer consists of discrete points—for each n, $\mathbf{ov}(\tau^n(T)) \cap \mathbf{ov}(\tau^{n+1}(T))$ is uncountable. See, for example, Figure 9, where $\lambda = 13/20$. To be more precise, if we rescale the initial triangle T so that the length of each of its sides is $b = \lambda/(1-\lambda)$, then for each n, the set $\mathbf{ov}(\tau^n(T))$ consists of 3^n disjoint triangles, each of side length λ^n.

(v) **The Closing of the Gap.** When $\lambda = 2/3$, there is still overlap with multiplicity, but there are no longer gaps at each iteration. This pattern continues for $\lambda > 2/3$. See Figure 9, which shows a λ value slighly less than 2/3 and Figure 10, which illustrates no gaps ($\lambda = 3/4$).

Summary of conclusions for 2D:

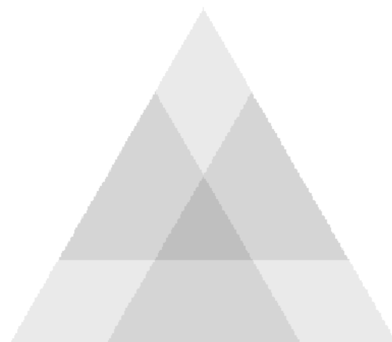


FIGURE 10. The first iteration of X_λ for $\lambda = \frac{3}{4}$. In this case, there are no gaps.

We have sketched features that come out differently for "IFS overlap-Sierpinski fractals," stressing differences that arise when passing from 1D examples where $\lambda \in (\frac{1}{2}, 1)$ to our analogous 2D attractors. Specifically, in the 2D case, for the range of values of λ, there are five separate cases for scaling numbers λ of interest, illustrating overlap features:

(i) $\lambda \in (1/2, (\sqrt{5}-1)/2)$: $\mathbf{ov}(\tau^n(T)) \cap \mathbf{ov}\tau^{n+1}(T)) = \varnothing$ (see Figure 7)
(ii) $\lambda = (\sqrt{5}-1)/2$: $\mathbf{ov}(\tau^n(T)) \cap \mathbf{ov}(\tau^{n+1}(T))$ consists of vertices (see Figure 8); simple overlap; central gaps
(iii) $\lambda \in (\sqrt{5}-1)/2, 2/3)$: $\mathbf{ov}(\tau^n(T)) \cap \mathbf{ov}(\tau^{n+1}(T))$ consists of 3^n disjoint triangles; overlap with multiplicity; central gaps (see Figure 9 for a Sierpinski figure whose gaps are very small)
(iv) $\lambda = 2/3$: central gaps close at $\lambda = 2/3$; overlap with multiplicity
(v) $\lambda \in (2/3, 1)$: no central gaps; overlap with multiplicity (see Figure 10).

6. Conclusions (the general case)

We now return to the general case of IFSs with essential overlap. In this case, the size of the overlap can nicely be expressed in terms of the column isometry from Definition 2.1. To recall the setting, we begin with the proof of Theorem 3.10 that was postponed.

Proof of Theorem 3.10. A system of measurable endomorphisms $\tau_1, \ldots, \tau_N$ in a finite measure space $(X, \mathcal{B}, \mu)$ is given, and it is assumed that $\mu(X) = 1$, and that

$$\mu = \frac{1}{N} \sum_{i=1}^{N} \mu \circ \tau_i^{-1}, \tag{6.1}$$

i.e., that μ is a (τ_i)-equilibrium measure. It then follows from Proposition 2.10 that the operators $F_i\colon f \mapsto \frac{1}{\sqrt{N}} f \circ \tau_i$ define a column isometry, i.e., that $\mathbb{F} = \begin{pmatrix} F_1 \\ \vdots \\ F_N \end{pmatrix}$ satisfies $\mathbb{F}^*\mathbb{F} = I_{L^2(\mu)}$.

Note that this identity spells out to

$$\sum_{i=1}^{N} F_i^* F_i = I_{L^2(\mu)}, \tag{6.2}$$

but that in general, the *individual* operators $F_i^* F_i$ are not projections. We have the lemma:

Lemma 6.1. *Let* $(\tau_i)_{i=1}^N$, μ, $\mathbb{F} = (F_i)_{i=1}^N$ *be as above, i.e.,* $\mathbb{F}$ *is a column isometry* $L^2(\mu) \to L^2(\mu)_N$. *Let* $\varphi_i := \frac{d\mu\circ\tau_i^{-1}}{d\mu}$. *Then*

$$F_i^* F_i = \frac{1}{N} M_{\varphi_i} \tag{6.3}$$

where M_{φ_i} *is the multiplication operator* $f \mapsto \varphi_i f$ *in* $L^2(\mu)$.

Proof. The result follows essentially from the argument in Lemma 2.8 above. By that argument we may pass to partitions of X. Let i be given, fixed; and, following Lemma 2.8, pass to a subset E in X such that there is a measurable mapping $\sigma_E\colon \tau_i(E) \to E$ with

$$\sigma_E \circ \tau_i|_E = \mathrm{id}_E; \tag{6.4}$$

see (2.13).

It follows that

$$F_i^* F_i f|_E = \frac{1}{N} \varphi_i f|_E$$

for all $f \in L^2(\mu)$. But the set E is part of a finite measurable partition of X, so the desired conclusion (6.3) holds on X. ☐

Proof of Theorem 3.10 *continued.* Our assertion is that $\mathbb{F}\mathbb{F}^* = \left(F_i F_j^*\right)_{i,j=1}^N$ is the identity operator in $L^2(\mu)_N$ if and only if the system is essential non-overlap.

Setting $\varphi_i := \frac{d\mu\circ\tau_i^{-1}}{d\mu}$ and using Lemma 2.8 we show that there are measurable and invertible point transformations $T_{i,j}\colon X \to X$, $i,j = 1,\dots,N$, such that $T_{i,i} = \mathrm{id}_X$, $1 \le i \le N$, and

$$F_i F_j^* = \frac{1}{N} (\varphi_j \circ \tau_i) T_{i,j}. \tag{6.5}$$

But, for each i, the function φ_i is supported on $\tau_i(X)$. So if

$$F_i F_j^* = \delta_{i,j} I_{L^2(\mu)}, \tag{6.6}$$

then

$$\varphi_i \equiv N \qquad \mu\text{-a.e. on } \tau_i(X)$$

and

$$\varphi_i \equiv 0 \qquad \mu\text{-a.e. on } X \setminus \tau_i(X) = \bigcup_{k\neq i} \tau_k(X).$$

The conclusion of the theorem is immediate from this; and we get the following corollary. □

Corollary 6.2. *Let the IFS $(\tau_i)_{i=1}^N$ be as specified in Theorem* 3.10 *above, and let $\mathbb{F}\colon L^2(\mu) \to L^2(\mu)_N$ be the corresponding column isometry, with $F_i\colon f \mapsto \frac{1}{\sqrt{N}} f \circ \tau_i$. Then $\mathbb{F}$ maps onto $L^2(\mu)_N$ if and only*

$$(F_i^* f) = \sqrt{N} \chi_{\tau_i(X)}(x) f(\sigma_i(x)) \qquad \mu\text{-a.e. } x \in X. \tag{6.7}$$

(Here the endomorphisms $\sigma_i\colon X \to X$ are specified in Lemma 2.8, and in particular $\sigma_i \circ \tau_i = \mathrm{id}_X$, $1 \le i \le N$.)

Theorem 6.3. *Let $N \in \mathbb{N}$, $N \geq 2$, be given, and let $(\tau_i)_{i \in \mathbb{Z}_N}$ be a contractive IFS in a complete metric space. let (X, μ) be the Hutchinson data; see Definition* 3.1. *Let P $(= P_{1/N})$ be the Bernoulli measure on $\Omega = \prod_1^\infty \mathbb{Z}_N = \mathbb{Z}_N^{\mathbb{N}}$; see Corollary* 3.5. *Let $\pi\colon \Omega \to X$ be the encoding mapping of Lemma* 3.3. *Set*

$$F_i f := \frac{1}{\sqrt{N}} f \circ \tau_i \qquad \text{for } f \in L^2(X, \mu) \tag{6.8}$$

and

$$S_i^* \psi := \frac{1}{\sqrt{N}} \psi \circ \sigma_i \qquad \text{for } \psi \in L^2(\Omega, P), \tag{6.9}$$

where σ_i denotes the shift map of (3.7).

(a) *Then the operator $V\colon L^2(X, \mu) \to L^2(\Omega, P)$ given by*

$$Vf = f \circ \pi \tag{6.10}$$

is isometric.

(b) *The following intertwining relations hold:*

$$VF_i = S_i^* V, \qquad i \in \mathbb{Z}_N. \tag{6.11}$$

(c) *The isometric extension $L^2(X, \mu) \hookrightarrow L^2(\Omega, P)$ of the (F_i)-relations is minimal in the sense that $L^2(\Omega, P)$ is the closure of*

$$\bigcup_n \bigcup_{i_1 i_2 \ldots i_n} S_{i_1} S_{i_2} \cdots S_{i_n} V L^2(X, \mu). \tag{6.12}$$

Proof. *Ad* (a)–(b): Let $f \in L^2(X, \mu)$, and let $\|\cdot\|_\mu$ and $\|\cdot\|_P$ denote the respective L^2-norms in $L^2(\mu)$ and $L^2(P)$. Then

$$\begin{aligned} \|Vf\|_P^2 &= \int_\Omega |f \circ \pi|^2 \, dP \\ &= \int_X |f|^2 \, d\left(P \circ \pi^{-1}\right) \\ &\underset{\text{by (3.13)}}{=} \int_X |f|^2 \, d\mu = \|f\|_\mu^2 . \end{aligned}$$

Moreover,

$$\begin{aligned} VF_i f &= (F_i f) \circ \pi \\ &= \frac{1}{\sqrt{N}} f \circ \tau_i \circ \pi \\ &\underset{\text{by (3.11)}}{=} \frac{1}{\sqrt{N}} f \circ \pi \circ \sigma_i \\ &= S_i^* V f, \end{aligned}$$

which is assertion (b).

Ad (c): Let $\psi \in L^2(\Omega, P)$, and let $\langle \cdot \mid \cdot \rangle_\mu$ and $\langle \cdot \mid \cdot \rangle_P$ denote the respective Hilbert inner products of $L^2(\mu)$ and $L^2(P)$. To show that the space in (6.12) is dense in $L^2(P)$, suppose

$$0 = \langle S_{i_1} \cdots S_{i_n} V f \mid \psi \rangle_P \tag{6.13}$$

for all n, all multi-indices $(i_1 \dots i_n)$, and all $f \in L^2(\mu)$. We will prove that then $\psi = 0$.

When $(i_1 \dots i_n)$ is fixed, we denote the cylinder set in Ω by

$$C(i_1, \dots, i_n) = \{ \omega \in \Omega \mid \omega_j = i_j,\ 1 \leq j \leq n \}. \tag{6.14}$$

Using now (6.7) in Corollary 6.2 on Ω, we get

$$S_{i_n}^* \cdots S_{i_1}^* \psi = N^{-n/2} \psi \circ \sigma_{i_1} \circ \cdots \circ \sigma_{i_n}.$$

Substitution into (6.13) yields

$$\int_\Omega \chi_{C(i_1,\dots,i_n)} \psi \, dP = 0.$$

We used the fact that (6.13) holds for all $f \in L^2(\mu)$. But the indicator functions $\chi_{C(i_1,\dots,i_n)}$ span a dense subspace in $L^2(\Omega, P)$ when n varies, and all finite words of length n are used. We conclude that $\psi = 0$, and therefore that the space in (6.12) is dense in $L^2(\Omega, P)$. $\square$

Remark 6.4. Note that by (6.11) the space in (6.12), part (c) of the theorem, is invariant under the operators S_i^*.

Acknowledgements. We are pleased to thank Dorin Dutkay for helpful conversations. We thank Brian Treadway for excellent typesetting, and for producing the graphics.

References

[Arv04] W. Arveson. The free cover of a row contraction. *Doc. Math.*, 9:137–161, 2004.

[Bar06] M. Barnsley. *SuperFractals*. Cambridge University Press, Cambridge, 2006.

[Bea91] A.F. Beardon. *Iteration of Rational Functions: Complex Analytic Dynamical Systems*, volume 132 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991.

[BHS05] M. Barnsley, J. Hutchinson, and Ö. Stenflo. A fractal valued random iteration algorithm and fractal hierarchy. *Fractals*, 13(2):111–146, 2005.

[BJMP04] L.W. Baggett, P.E.T. Jorgensen, K.D. Merrill, and J.A. Packer. An analogue of Bratteli-Jorgensen loop group actions for GMRA's. In C. Heil, P.E.T. Jorgensen, and D.R. Larson, editors, *Wavelets, Frames, and Operator Theory (Focused Research Group Workshop, College Park, Maryland, January 15–21, 2003)*, volume 345 of *Contemp. Math.*, pages 11–25. American Mathematical Society, Providence, 2004.

[BJMP05] L.W. Baggett, P.E.T. Jorgensen, K.D. Merrill, and J.A. Packer. Construction of Parseval wavelets from redundant filter systems. *J. Math. Phys.*, 46(8):083502, 28 pp., 2005. doi:10.1063/1.1982768.

[BJMP06] L.W. Baggett, P.E.T. Jorgensen, K.D. Merrill, and J.A. Packer. A non-MRA C^r frame wavelet with rapid decay. *Acta Appl. Math.*, 89:251–270, 2006. doi:10.1007/s10440-005-9011-4.

[BV05] J.A. Ball and V. Vinnikov. Lax–Phillips Scattering and Conservative Linear Systems: A Cuntz-Algebra Multidimensional Setting. *Mem. Amer. Math. Soc.*, 178(837), 2005.

[CF05] R.E. Curto and L.A. Fialkow. Solution of the truncated hyperbolic moment problem. *Integral Equations Operator Theory*, 52(2):181–218, 2005.

[CI06] E. Christensen and C. Ivan. Spectral triples for AF C^*-algebras and metrics on the Cantor set. *J. Operator Theory*, 56:17–46, 2006.

[CK80] J. Cuntz and W. Krieger. A class of C^*-algebras and topological Markov chains. *Invent. Math.*, 56(3):251–268, 1980.

[Cun77] J. Cuntz. Simple C^*-algebras generated by isometries. *Comm. Math. Phys.*, 57:173–185, 1977.

[DJ05] D.E. Dutkay and P.E.T. Jorgensen. Hilbert spaces of martingales supporting certain substitution-dynamical systems. *Conform. Geom. Dyn.*, 9:24–45, 2005. http://www.ams.org/ecgd/2005-09-02/S1088-4173-05-00135-9/home.html.

[DJ06a] D.E. Dutkay and P.E.T. Jorgensen. Hilbert spaces built on a similarity and on dynamical renormalization. *J. Math. Phys.*, 47(5), 2006.

[DJ06b] D.E. Dutkay and P.E.T. Jorgensen. Iterated function systems, Ruelle operators, and invariant projective measures. *Math. Comp.*, 75:1931–1970, 2006.

[DJ06c] D.E. Dutkay and P.E.T. Jorgensen. Methods from multiscale theory and wavelets applied to non-linear dynamics. In D. Alpay, editor, *Wavelets, Multiscale Systems and Hypercomplex Analysis*, volume 167 of *Oper. Theory Adv. Appl.*, pages 87–126. Birkhäuser, Boston, 2006. http://arxiv.org/abs/math.DS/0411371.

[DJ06d] D.E. Dutkay and P.E.T. Jorgensen. Wavelets on fractals. *Rev. Mat. Iberoamericana*, 22(1):131–180, 2006.

[Erd40] P. Erdős. On the smoothness properties of a family of Bernoulli convolutions. *Amer. J. Math.*, 62:180–186, 1940.

[FLP94] L. Flatto, J.C. Lagarias, and B. Poonen. The zeta function of the beta transformation. *Ergodic Theory Dynam. Systems*, 14(2):237–266, 1994.

[GRS01] M. Gibbons, A. Raj, and R.S. Strichartz. The finite element method on the Sierpinski gasket. *Constr. Approx.*, 17(4):561–588, 2001.

[HR00] J.E. Hutchinson and L. Rüschendorf. Random fractals and probability metrics. *Adv. in Appl. Probab.*, 32(4):925–947, 2000.

[Hut81] J.E. Hutchinson. Fractals and self similarity. *Indiana Univ. Math. J.*, 30:713–747, 1981.

[JK03] P.E.T. Jorgensen and D.W. Kribs. Wavelet representations and Fock space on positive matrices. *J. Funct. Anal.*, 197:526–559, 2003.

[Jor04] P.E.T. Jorgensen. Iterated function systems, representations, and Hilbert space. *Internat. J. Math.*, 15:813–832, 2004.

[Jor05] P.E.T. Jorgensen. Measures in wavelet decompositions. *Adv. Appl. Math.*, 34:561–590, 2005.

[Jor06] P.E.T. Jorgensen. *Analysis and Probability: Wavelets, Signals, Fractals*, volume 234 of *Grad. Texts in Math.* Springer, New York, 2006.

[JP96] P.E.T. Jorgensen and S. Pedersen. Harmonic analysis of fractal measures. *Constr. Approx.*, 12:1–30, 1996.

[JP98] P.E.T. Jorgensen and S. Pedersen. Dense analytic subspaces in fractal L^2-spaces. *J. Analyse Math.*, 75:185–228, 1998.

[Kaw05] K. Kawamura. The Perron-Frobenius operators, invariant measures and representations of the Cuntz-Krieger algebras. *J. Math. Phys.*, 46(8):083514, 6, 2005.

[KLPL06] David W. Kribs, Raymond Laflamme, David Poulin, and Maia Lesosky. Operator quantum error correction. *Quantum Inf. Comput.*, 6(4-5):382–398, 2006.

[Kol77] A.N. Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Springer-Verlag, Berlin–New York, 1977. reprint of the 1933 original; English translation: *Foundations of the Theory of Probability*, Chelsea, 1950.

[Kri05] D.W. Kribs. A quantum computing primer for operator theorists. *Linear Algebra Appl.*, 400:147–167, 2005.

[KS93] R. Kumar and S.D. Sharma. Inner functions and substitution operators. *Acta Sci. Math. (Szeged)*, 58(1-4):509–516, 1993.

[Kwa04] J. Kwapisz. Transfer operator, topological entropy and maximal measure for cocyclic subshifts. *Ergodic Theory Dynam. Systems*, 24(4):1173–1197, 2004.

[LS05] E.H. Lieb and R. Seiringer. Stronger subadditivity of entropy. *Phys. Rev. A (3)*, 71(6):062329, 9 pp., 2005.

[LW00] J.C. Lagarias and Y. Wang. Orthogonality criteria for compactly supported refinable functions and refinable function vectors. *J. Fourier Anal. Appl.*, 6(2):153–170, 2000.

[Pop89] G. Popescu. Isometric dilations for infinite sequences of non-commuting operators. *Trans. Amer. Math. Soc.*, 316:523–536, 1989.

[PS96] Yuval Peres and Boris Solomyak. Absolute continuity of Bernoulli convolutions, a simple proof. *Math. Res. Lett.*, 3(2):231–239, 1996.

[Sin05] R.K. Singh. A survey of weighted substitution operators and generalizations of Banach-Stone theorem. *Int. J. Math. Math. Sci.*, (6):937–948, 2005.

[Sol95] B.M. Solomyak. On the random series $\sum \pm\lambda^n$ (an Erdős problem). *Ann. of Math. (2)*, 142:611–625, 1995.

[Sol98] B.M. Solomyak. Non-linear iterated function systems with overlaps. *Period. Math. Hungar.*, 37(1-3):127–141, 1998. International Conference on Dimension and Dynamics (Miskolc, 1998).

[Str00] R.S. Strichartz. Mock Fourier series and transforms associated with certain Cantor measures. *J. Anal. Math.*, 81:209–238, 2000.

[Str06] Robert S. Strichartz. *Differential equations on fractals*. Princeton University Press, Princeton, NJ, 2006. A tutorial.

[SU00] R.S. Strichartz and M. Usher. Splines on fractals. *Math. Proc. Cambridge Philos. Soc.*, 129(2):331–360, 2000.

[SV98] N. Sidorov and A. Vershik. Ergodic properties of the Erdős measure, the entropy of the golden shift, and related problems. *Monatsh. Math.*, 126(3):215–261, 1998.

(Palle E. T. Jorgensen) Department of Mathematics, The University of Iowa, Iowa City, IA 52242-1419, U.S.A.

E-mail address: `jorgen@math.uiowa.edu`

URL: `http://www.math.uiowa.edu/~jorgen/`

(Keri Kornelson, Karen Shuman) Department of Mathematics and Statistics, Grinnell College, Grinnell, IA 50112-1690, U.S.A.

URL: `http://www.math.grinnell.edu/~kornelso/`

URL: `http://www.math.grinnell.edu/~shumank/`